REPORT RESUMES

ED 002 882 24

A STUDY OF THE RELATIONSHIP BETWEEN OBSERVED CLASSROOM BEHAVIORS OF ELEMENTARY STUDENT TEACHERS, PREDICTORS OF THOSE BEHAVIORS, AND RATINGS BY SUPERVISORS.

BY- WILK, ROGER E. AND OTHERS

MINNESOTA UNIV., MINNEAPOLIS, COLL. OF EDUCATION

REPORT NUMBER CRP-473 PUB DATE JAN 62

EDRS PRICE MF-$0.18 HC-$4.76 119P.

DESCRIPTORS- *STUDENT TEACHERS, PRESERVICE EDUCATION, BEHAVIOR PATTERNS, *CLASSROOM COMMUNICATION, ELEMENTARY SCHOOL TEACHERS, PREDICTIVE MEASUREMENT, TEACHER SUPERVISION, TEACHER EDUCATION, TEACHER PLACEMENT, *PLACEMENT, *ADMISSION (SCHOOL), *SUPERVISION, CLASSROOM ENVIRONMENT, OBSERVATION, MINNEAPOLIS, MINNESOTA

THREE PARTICULAR PROBLEMS WERE THE CONCERN OF THIS STUDY--(1) THE VALIDITY OF ADMISSIONS DATA USED IN ELEMENTARY TEACHER EDUCATION PROGRAMS, (2) THE EFFECTS OF DIFFERENT STUDENT TEACHING PLACEMENTS, AND (3) THE RELATIONSHIP BETWEEN JUDGMENT RATINGS OF STUDENT TEACHERS MADE BY STUDENT TEACHING SUPERVISORS USING THE REGULARLY EMPLOYED OBSERVATIONAL PROCEDURES OF THEIR OFFICE AND COMPARABLE RATINGS MADE BY RESEARCH OBSERVERS USING PROCEDURES ESPECIALLY SET UP FOR THE PURPOSE OF THIS STUDY. THE DESIGN TREATED FOUR SOURCES OF VARIATION AS EXPERIMENTAL VARIABLES--(1) STUDENT TEACHERS' GRADE LEVEL PREFERENCE, (2) STUDENT TEACHERS' QUARTER OF EXPERIENCE, (3) CLASSROOM GRADE LEVEL, AND (4) ORDER OF PREFERENCE IN TWO DIFFERENT GRADE LEVEL ASSIGNMENTS. THE CONTROLLED VARIABLES (THOSE TREATED IN SUCH A WAY THAT THEY WOULD NOT BIAS THE OBSERVATIONS) WERE--(1) THE PROFESSIONAL EDUCATION OF THE STUDENT TEACHER, (2) THE SEX OF THE STUDENT TEACHER, (3) TRAVEL BY THE STUDENT TEACHER, AND (4) SOCIOECONOMIC LEVEL OF THE CLASSROOM. ON THE BASES OF STUDENT TEACHER OBSERVATIONS AND INTERVIEWS, AND SPECIFIC TEST ADMINISTRATIONS, THE FOLLOWING CONCLUSIONS WERE REACHED--(1) ADMISSION DATA WERE NOT RELIABLE FOR CLASSROOM BEHAVIOR PREDICTIONS, (2) CLASSROOM PLACEMENTS APPEARED TO BE MORE CLOSELY RELATED TO SPECIFIC STUDENT TEACHER RESPONSES TO INDIVIDUAL PUPILS AND TO A CLASS THAN THEY WERE TO GENERALIZED TENDENCIES TO USE CERTAIN CATEGORIES OF VERBAL BEHAVIOR, AND (3) RATINGS BETWEEN THE SUPERVISORS AND THE RESEARCH OBSERVERS ON EITHER PERSONALITY OR CLASSROOM PERFORMANCE JUDGMENTS GENERALLY EXHIBITED A HIGH DEGREE OF AGREEMENT. (JH)

ED 002552

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

A STUDY OF THE RELATIONSHIP BETWEEN OBSERVED CLASSROOM BEHAVIORS OF ELEMENTARY STUDENT TEACHERS, PREDICTORS OF THOSE BEHAVIORS, AND RATINGS BY SUPERVISORS.

by
Roger E. Wilk and William H. Edson

with
Don Davies and Naomi C. Chase

College of Education
University of Minnesota

The research reported herein was performed pursuant to a contract with the United States Office of Education, Department of Health, Education, and Welfare.

January, 1962

3414A

Acknowledments

Acknowledgment is due first of all to the Cooperative Research Program of the U. S. Office of Education for the financial support of this study. The research reported herein was performed pursuant to a contract with the United States Office of Education, Department of Health, Education & Welfare.

An undertaking with the depth and scope of the student teaching study described in this report required a dedication to the cause of improving teacher education. Such consideration of the task was found in the participants of all phases of the study. The College of Education personnel who accepted the administrative responsibility for the project are grateful to those who so willingly participated from the outset to the completion of the study.

We are particularly indebted to the Supervising Teachers of Minneapolis and St. Paul who accepted with understanding the numerous visits made for the purpose of collecting data. To the Administrative Staffs of these Twin Cities we extend thanks for opening their schools to our quest for Supervising Teachers willing to join us in the study.

The elementary women students who were the "sample" of seniors to be observed should also receive considerable recognition. In addition to the regular visits arranged by their University Supervisors, these students were called upon to welcome, in addition, over twice as many more observations by the research observers. The students were assigned to schools according to the design required for statistical treatment of data collected. Although the randomizing was done within geographic areas not remote from student residences, there were a few travel hardships. These were accepted with gradifying cooperation.

The statistical design of the study and the method of computing the reliability of observers were developed with the generous assistance of Dr. Raymond O. Collier, Jr., Dr. Robert L. Lathrop, and Mr. Donald G. MacEachern, all members of the faculty of the College of Education.

Although they were fully constituted staff members for the project, hearty thanks is extended to the research staff for their congenial good work.

In order to prevent loss of individuality among the members of the large group of essential participants, their names follow in total.

Research Staff

Research Coordinator
Edward L. Kelly

Research Observers
Jane R. Collins
Donald L. Johnson
Donald L. Lantz
Cornelia D. McCune
H. Bradley Sagen

University Student Teaching Supervisors
Karl R. Anselm
Grant Bateman
William R. Harmer
Ruth D. Johnson
Inga B. Kromann
Katherine M. O'Donnell
Alice A. Tuseth

Supervising Teachers

Minneapolis Public Schools
Arthur J. Lewis, Assistant Superintendent, Elementary Education

Supervising Teacher	School	Supervising Teacher	School
Mr. Bradley Bentson	Warrington	Mrs. Janet James	Whittier
Mrs. Agnes Berntson	Motley	Mrs. Jessamine Johnson	Keewaydin
Mrs. Lenore Brightbill	Irving	Mrs. Kathleen Kieffer	Keewaydin
Mr. Frank Bruce	Longfellow	Miss Alice Lang	Madison
Mrs. Ella Carlson	Longfellow	Mrs. JoAnn Lange	Lowell
Miss Dorthea Donohue	Blaine	Mrs. Elizabeth Long	Minnehaha
Mr. Donald Eblen	Schiller	Miss Harriette Magnuson	Whittier
Mrs. Dorothy Engels	Grant	Miss Eulalia Murphy	Hay
Mr. Douglas Farrell	Longfellow	Miss Leona Saatoff	Loring
Mrs. Carol Freeburg	Grant	Mrs. Florence Sandvik	Windom
Mrs. Bessie Gullickson	Pratt	Mr. William Scott	Windom
Mrs. Caroline Gustafson	Ericsson	Mrs. Hazel Toda	Motley
Mr. Melvin Hoaglund	McKinley	Miss Margaret Ulrich	Fulton
Miss Alice Hurley	Northrop	Mr. Donald Wagner	Pratt
Mrs. Lois Jackson	Lind		

St. Paul Public Schools
Richard G. Hansen, Assistant Superintendent, Elementary Education

Supervising Teacher	School	Supervising Teacher	School
Mrs. Ione Berglund	Jefferson	Mrs. Hedi Koehler	St. Anthony
Miss Audrey Bodelson	Lincoln	Mrs. Edna Mudge	Desnoyer Pk.
Miss Marilyn Campbell	Rice	Mrs. Phyllis Rupp	Highland Pk.
Mr. Robert Grant	Linwood	Mrs. Helen Ryan	Lincoln
Miss Sara Ingebrigtson	St. Anthony		

CONTENTS

Acknowledgments iii

List of Tables vii

Chapter I. The Problem, Its Development and Setting 1
  Introduction 1
  The Genesis of the Study 2
  Questions to be Studied 2
  Conceptual Framework 4
  The Elementary Education Program 6

Chapter II. The Design and Conduct of the Study 9
  Introduction 9
  The Design of the Study 9
  The Sources of Variability Studied 9
  The Cross-over Design 12
  Observation and Measurement of Variables 13
  Classifying Classroom Verbal Behaviors 14
  The Conduct of the Study 18
  Planning with the Participants 18
  Training the Observers 19
  Assignment of Student Teachers 20
  Conducting the Observations 21

Chapter III. The Observed Behaviors of Student Teachers 23
  Data from the OScAR III 23
  Scoring and Scaling 23
  Student Teacher Activities 25
  Subject Areas Taught 26
  Variety of Materials Used 26
  Pupil Response 27
  Relation of Student Teacher Activities to Pupil Response 28
  Data from the Classification of Verbal Behaviors 29
  Observer Reliability 29
  Weighted and Unweighted Scoring 33

Chapter IV. The Relations of Observed Behaviors to Classroom Placement Variables 36
  The Cross-over Design 36
  Analysis of OScAR III Scales 38
  Analysis of Verbal Behaviors 45

Chapter V. The Relation of Predictor Variables to Observed Behaviors 50
  Relationships Among the Variables 50
  Multiple Regression Analysis 51
  Predicting n Dom 53
  Predicting n Intg 54

CONTENTS con't

Chapter VI. The Comparisons of Ratings by Observers and Supervisors 55
    Ratings of Personality Variables 55
    Ratings of Teaching Behaviors 57

Chapter VII. Summary and Conclusions 63
    The Subject of the Study 63
    The Design of the Study 64
    The Sources of Variability 64
    The Cross-over Design 65
    Observation and Measurement of Variables 66
    The Conduct of the Study 67
    The Observed Behaviors of Student Teachers 68
    The OScAR III 68
    The Minnesota System of Interaction Analysis 69
    The Relation of Observed Behaviors to Classroom Placement Variables 69
    Analysis of the OScAR III Scales 70
    Analysis of Verbal Behavior 71
    The Comparison of Observed Behaviors and Ratings by Superviisors 72
    The Relation of Predictor Variables to Observed Behaviors 73
    Conclusions 74

References 77

Appendix A. A-1 - Socio-economic Data
    A-2 - University Supervisors' Rating Forms
    A-3 - The OScAR III
    A-4 - The Minnesota System of Categorizing Verbal Behaviors

Appendix B. Frequency of Occurrence of Classroom Behaviors Recorded in Ten Observations of Thirty-six Elementary Student Teachers Using the Categories of the OScAR III.

Appendix C. C-1 - Regression analysis predicting n Dom (RA) scores.
    C-2 - Regression analysis predicting n Intg (RA) scores.

LIST OF TABLES

Table	Title	Page
2.1	Categories for the Classification of Teacher Verbal Behaviors (From the Minnesota System of Interaction Analysis, 1958)	15
3.1	Item Numbers, Means and Standard Deviations for Eight Scales Derived from the OScAR III	24
3.2	Intercorrelations of OScAR III Scales Describing Classroom Activities and Pupil Response	27
3.3	The Partition of the Sum of Squares for Determining the Reliability of an Average Judge When Several Judges Observe One Teacher at Different Time Periods	30
3.4	The Partition of the Sum of Squares for Determining the Reliability of an Average Judge When Several Judges Observe One Teacher at Different Time Periods Replicated for Several Teachers	31
3.5	The Reliabilities for the Average Judge and For a Composite of Five Judges Estimated from a Series of Thirteen Observations Over a Two Quarter Period for n Intg and n Dom Variables	30
3.6	Mean *A Priori* and Reciprocal Average Weights for each Category of Verbal Behavior Comprising n Dom and n Intg Scores	33
3.7	Mean, Standard Deviation of Unweighted and Reciprocal Average Scores and Their Correlation for n Dom and n Intg	34
4.1	Analysis of Variance Model for Counterbalanced Observations of Student Teachers at Both Upper and Lower Grades over a Two Quarter Period Stratified According to Grade Level Preference	36
4.1a	Computation Formulas for Sums of Squares for Analysis of Variance Model	37
4.2	Analysis of Variance of Teacher-Individual Activities for Counterbalanced Observations of Student Teachers at Both Upper and Lower Elementary Grades over a Two Quarter Period Stratified According to Grade Level Preference	38
4.3	Analysis of Variance of Teacher Class Activities for Counterbalanced Observations of Student Teachers at Both Upper and Lower Elementary Grades over a Two Quarter Period Stratified According to Grade Level Preference	39
4.4	Analysis of Variance of Teacher Group Activities for Counterbalanced Observations of Student Teachers at Both Upper and Lower Elementary Grades over a Two Quarter Period Stratified According to Grade Level Preference	40

LIST OF TABLES con't

Table	Title	Page
4.5	Analysis of Variance of Pupil Class Activities for Counterbalanced Observations of Student Teachers at Both Upper and Lower Elementary Grades over a Two Quarter Period Stratified According to Grade Level Preference	40
4.6	Analysis of Variance of Pupil Disruptive Behavior for Counterbalanced Observations of Student Teachers at Both Upper and Lower Elementary Grades over a Two Quarter Period Stratified According to Grade Level Preference	41
4.7	Analysis of Variance of Classroom, Variety of Materials for Counterbalanced Observations of Students Teachers at Both Upper and Lower Elementary Grades over a Two Quarter Period Stratified According to Grade Level Preference	41
4.8	Analysis of Variance of Classroom Emotional Climate for Counterbalanced Observations of Student Teachers at Both Upper and Lower Elementary Grades over a Two Quarter Period Stratified According to Grade Level Preference	43
4.9	Analysis of Variance of Classroom Verbal Emphasis for Counterbalanced Observations of Student Teachers at Both Upper and Lower Elementary Grades over a Two Quarter Period Stratified According to Grade Level Preference	43
4.10	Analysis of Variance of Teacher Verbal Output for Counterbalanced Observations of Student Teachers at Both Upper and Lower Elementary Grades over a Two Quarter Period Stratified According to Grade Level Preference	44
4.11	Analysis of Variance of Need Integration Scroes for Counterbalanced Observations of Student Teachers at Both Upper and Lower Elementary Grades over a Two Quarter Period Stratified According to Grade Level Preference	45
4.12	Analysis of Variance of Need Dominance Scores for Counterbalanced Observations of Student Teachers at Both Upper and Lower Elementary Grades over a Two Quarter Period Stratified According to Grade Level Preference	45
4.13	Analysis of Variance of Need Integration (RA) Scores for Counterbalanced Observations of Student Teachers at Both Upper and Lower Elementary Grades over a Two Quarter Period Stratified According to Grade Level Preference	47
4.14	Analysis of Variance of Need Dominance (RA) Scores for Counterbalanced Observations of Student Teachers at Both Upper and Lower Elementary Grades over a Two Quarter Period Stratified According to Grade Level Preference.	47

LIST OF TABLES con't

Table	Title	Page
5.1	Intercorrelations Among Predictor Variables and Among Predictor and Criterion Variables	49
6.1	The Correlation Between n Dom and n Intg Scores Obtained from Judgments of University Supervisors and Those Derived from Procedures Used by Research Observers	55
6.2	Data on Supervisors' and Observers' Ratings of Student Teachers on Four Variables Giving Mean, Standard Deviation, Reliabilities, the Correlation Between Ratings, and the Correlation Corrected for the Unreliability of Supervisors' and Observers' Ratings	57
6.3	Distribution by Supervisor and Analysis of Variance of Z Score Discrepancies (Supervisor's Mean Minus Counselors' Mean) for Creativity Scores	59
6.4	Distribution by Supervisor and Analysis of Variance of Z Score Discrepancies (Supervisor's Mean Minus Counselors' Mean) for Interpersonal Relations Scores	59
6.5	Distribution by Supervisor and Analysis of Variance of Z Score Discrepancies (Supervisor's Mean Minus Counselors' Mean) for Success as Student Teacher Ratings	60

Chapter I

The Problem, Its Development and Setting

Introduction

Are appraisals of teacher education applicants made at the time of admission related to descriptions of their subsequent classroom behavior? This question stimulated the design of the investigation. Its relevance lies in its relationship to the broader problem of the selection and retention of students in teacher education.

Through its program of selection and retention each school or college controls the quality of the teachers it provides for the schools. Selection and retention policies derived from the purposes of the institution are placed into operation by the decisions that are made about individual students. Before a policy may be applied the college must obtain information about an applicant. Since collecting and interpreting information about individual college students is, or ought to be, a function for the student personnel worker, he is especially concerned about the decision making procedures.

Specific problems arise for the student personnel worker as he participates in the decision making process. What behaviors are relevant to the decision? What data are to be collected? What observations should be made? How shall interpretations be given, and how shall predictions be made? To seek answers to such questions research must be an integral part of the student personnel program. The search for answers to questions such as those posed is made more difficult by the lack of an adequate criterion of teaching success. Those who would evaluate teachers have not yet been able to agree upon what is meant by teacher effectiveness nor to suggest a methodology for its determination. Although much of educational research is fettered by the criterion problem, there are areas which can and need to be studied.

Teacher education needs to know more about its entering and graduating students and what happens to them in the intervening years. Methods of observation need to be developed and improved. Psychometric variables need to be studied to understand their interrelationships. Such research should be designed to lead to precise descriptions and predictions of individual behavior and performance. Although these studies may contribute little to the question of determining teacher effectiveness, they will enable schools and colleges to know their students and the impact of their selection and retention procedures, as well as their program of study.

The preceding paragraphs represent the position of the Minnesota Studies of Student Personnel Work in Teacher Education. To find ways to improve the functions of teacher education student personnel work is a primary aim of this series of studies of which the present study is one. This series recognizes the importance of adequate description and prediction, especially, since no acceptable criterion for validity is available.

With this general introduction to the point of view of the Minnesota Studies, the remainder of this chapter will be devoted to a presentation of the specific questions to be studied, their genesis, the conceptual framework of the study, and the setting in which the questions were asked. Subsequent chapters will deal with the design of the study, present results and discuss the implications.

The Genesis of the Study

In January, 1957, the Student Personnel Office of the College of Education at the University of Minnesota began work on a research contract with the United States Office of Education under its Cooperative Research Program to appraise the use of the interview as a means of screening candidates for teacher education. Although the experiment investigated several related questions concerning the interview, the principle concern was whether or not trained counselors using interview information could make reliable judgments of teacher education applicants. Within the conceptual framework of the study, results showed that, on the whole, counselors can make reliable judgments from the interview, but that some judgments are more reliable than others. (Edson and Wilk, 1958). (The specific variables that were used appear later in this chapter.) It is certainly important to know that there are not large differences among the judgments made by the different counselors under the conditions of this study. Such consistency suggests that all counselors employed a similar basis for their judgments. This consistency is a desirable condition for any further developments of the use of the interview as an assessment device.

The very important question of the relationship of counselor judgment to subsequent behavior of students was outside the scope of the interview study. Until further investigation could ascertain whether or not the type of judgments made in the interview study were indicative of future performance, these judgments were of limited use and their meaning only speculative. The major impetus of the present study was a concern for finding the predictive value of admission data including the interview judgments.

The staff of the Student Personnel Office depended on the counsel and the cooperation of other college departments as they considered the type of study that would assess student performance and relate the assessments to admission data. Through the conduct of its normal functions, the office maintains a close communication with other phases of the college program and enjoys a helpful, reciprocal relationship. From the inception of the study, the staff and the Elementary Education staff were essential participants in this project. Because the most convenient observations of performance could be made during student teaching and because of a number of questions of interest to the Student Teaching Office could be incorporated into a study, the Student Teaching Office joined the Personnel Office in this project.

Questions to be Studied

The discussion in the previous section indicates that this study developed from an interest in seeking out the behavioral correlates of admission data, including interview judgments. The first general question which this study sought to answer may be written as follows:

> What is the relationship between the observed behaviors of student teachers and data collected earlier at the time of admission to the professional courses in teacher education?

The report of the interview study presents a description of the admission program of the College of Education at the University of Minnesota and its relation to the professional sequence of courses (Edson and Wilk, 1958, p. 2). In brief, students seeking admission to the professional courses in teacher education offered only during the junior and senior year must make application and take a battery of psychological tests, a speech and hearing test, a health examination, and complete admissions interviews with a counselor in the Student Personnel Office and an advisor in the teaching field. The first general question seeks to find whether data gathered at the beginning of the junior year can be predictive of performance in student teaching in the senior year.

Plans for the study recognized the necessity for controlling or estimating a number of variables related to student teaching. The sources of variation that were of concern are included in a second general question:

> What is the relationship of student teaching preference, experience, and the classroom situation to observed behavior of student teachers?

The design of the study provided for the identification of the situational and experiencial influences that affected the observations associated with the assessment of student teacher performance. It was desirable to be able to attribute observed variability to the individual differences of the student teachers, separating out the other effects.

The third general question concerns the assessment of the classroom behavior variables. Crucial to the study were answers to the questions about who should be observers and what type of assessments should be made. The observers selected were persons deemed to be sensitive to the dynamics of personal behavior. Teaching experience in the area to be observed was not a consideration in their selection. This decision was based on the fact that the variables used in the interview study, which were also necessary for investigating the first general question, did not require the assessment of pedogogical variables.

Consideration of the problem of selecting the methods of observation to be used and the choice of observers led to a statement of the third question in these two ways.

> (1) Using the personality variables of concern in this study as the basis for comparison, what is the relationship of ratings of student teachers made by experienced student teaching supervisors using their established observational methods and the ratings derived from the observational techniques designed for this study and employed by research observers?

(2) Using certain teaching behaviors as the basis for comparison, what is the relationship of ratings of student teachers by student teaching supervisors using their established observational methods and the ratings made by research observers using the methods designed for this study?

The final question concerned the more general problem of describing the activities of the classroom and was asked in two ways.

What are the activities of the student teacher in the classroom?

How do pupils respond to student teachers?

Conceptual Framework

The present study was conceived within the framework of the Minnesota Studies for Student Personnel Work in Teacher Education. Some further discussion of these studies is given at this point as a basis for understanding the project that is the subject of this report. The main stream of the Minnesota studies is a projected, twelve year longitudinal study of students' persistence in teacher education and teaching. This twelve year study is descriptive in nature and seeks to identify the characteristics of students who persist in the various curriculums of teacher education at Minnesota. The method employed is empirical, for the purposes are to describe, to discover relationships, and to generate hypotheses. This longitudinal study will analyze data which are normally gathered as part of the student personnel program. Essential to the research is separate consideration of each teacher education curriculum and the acceptance of persistence as the criterion. Although persistence in teacher education falls short of the ideal criterion of teaching effectiveness, it has the advantage of being operationally defined, and it does have practical significance for teacher education institutions.

In addition to a search for adequate descriptions, the Minnesota studies include projects that are more experimental in method and that look to theoretical constructs as the source of scales of measurement and hypotheses for tests. The present study and its predecessor, the interview study, are of this type. These studies adopt a view of personality that employs "needs" constructs in formulating hypotheses concerning individual behavior. Drawing primarily from the work of Murray (1938) and H. H. Anderson (1939), the interview study stated two propositions (Edson and Wilk, 1958, p. 5). The first postulated that persistence in teacher education required at least two qualities, Achievement potential and need Achievement. The second proposed that interest in the field of teaching was related to the variables of need Dominance, need Integration, plus "Security." It is the first two of the latter set of three variables that were important to the study reported here.

The formulation presented in the interview study suggested that appropriately prepared persons entering the teaching field would be motivated primarily either by a need to work with other people or a desire for the security that the teaching profession could bring. The need to work with others might stem either from a need to dominate them (need Dominance) or a need to be a socially integrative factor in their growth and development (need Integration). To provide a more refined definition, the following descriptions of the behavioral characteristics of these variables were developed in the earlier study.

n Dom (need Dominance)
- Sets definite structure and organization
- Sets high standards
- Provides rules, regulations, procedures
- Interaction with pupils often characterized by
    1. Persuasion, prohibition, commands
    2. Blame, criticism, sarcasm
    3. Threat, withholding approval
    4. Lack of regard for students' feelings
    5. Intolerance of differences

n Intg (need Integration)
- Sets broad structure and organization
- Provides for differentiation of tasks
- Flexibility of standards
- Interaction with pupils characterized by
    1. Support, encouragement, praise
    2. Freedom of student expression
    3. Attention to students' feelings
    4. Encouragement of different methods of approach
    5. Promotion of acceptance of others

These descriptions characterize the highest scale values on two separate scales of measurement. They were not considered as bi-polar variables representing opposite ends of a continuum, for such a representation would have precluded the possibility that any person could evidence some degree of both variables. Rather, they were considered as separate qualities which could be assessed independently and which were moderately negatively correlated. In the interview study, counselors judged the degree to which the ends of these variables were present as evidenced from admission data. Thus, any given person might have low or moderate scores on both scales, a high score on one and low on the other, but probably not high scores on both scales. Data from the interview study showed that the scales correlated at the -.30 level when counselors' judgments were based on the interview and test score data. With this summary, the relation of these variables to the present study can be explored.

The definitions of these variables previously outlined in this section make a number of the relevant classroom behaviors explicit. Both Anderson (1939) and Flanders (1960) have described similar behaviors although each in a somewhat different context, and their contribution to this study is recognized. In general, however, a high degree of n Dom would be evidenced by the teacher's exerting direct influence on the behavior of the pupils. Teacher behavior is said to have "direct influence" when its effect is to restrict the pupils' freedom of action. A high degree of n Intg would cause the teacher to influence pupils' behaviors indirectly, permitting more freedom of action on the part of the pupil and encouraging him to expand on his own ideas, to determine the way in which he might respond. It seems consistent with the definition of the interview study that teachers having high n Dom would use direct attempts to influence pupils toward specific, designated goals. It also seems logical that teachers with high n Intg would exhibit indirect influence in attempting to guide pupils toward differentiated, individualized goals.

As conceived in this study, there is no pedagogical value attached to either of these types of behavior. Conceivably, both direct and indirect influence attempts would be appropriate in different classroom circumstances. As outlined here, however, it is implied that teachers with differing degrees of these needs will, in general, evidence differences in their verbal behavior, as reflected in their attempts to influence pupil behavior. As stated in the previous section, the first question for study concerns the relation of admissions data to subsequent behavior in teacher education. The first step in this evaluation of counselors' judgments and other admissions data was to determine a relevant behavioral criterion against which a predictive relationship could be established. The first question then became:

How can n Dom and n Intg be observed?

Student teaching was accepted as the setting in teacher education in which the interpersonal manifestations of these variables are most likely to be evident. The sine qua non of the teaching process is the communication between teacher and pupils. Smith (1960) provides an excellent statement of this fact. Because the classroom communications are essential to the teaching process, they provide prominent phenomena observable among all teachers. If the personality needs described above do motivate entrance into teaching, and if these needs do influence the nature of the teachers' behavior toward children, the effect will be observable in the teachers' verbal behaviors.

The purpose of this argument has been to point out the relevance of teachers' verbal behaviors as a criterion for testing the validity of counselors' judgments. The rationale has suggested that the definitions which serve as the basis for counselor judgments could be directly observable in the verbal behavior of the teacher in the teaching situation. In general, the manner in which teachers attempted to influence pupils' behaviors, direct or indirect, are logically related to the original definitions of n Dom and n Intg.

Determining whether the criterion could be reliably determined constituted the second step in the validating process. The answer to this question depended upon the availability of the method of categorizing teachers' verbal behaviors, and upon training observers to make observations with a high degree of reliability. These problems and their solutions are presented in the following chapter, which deals with the design and the conduct of the study. The final step in the validation process, determining the predictive relationship, is presented in the results of this study.

The design and conduct of the study must have specific reference to the situation in which the hypotheses are put to a test. For this reason the characteristics of the setting are presented next.

The Elementary Education Program

The interview study to which reference has been made was conducted with women students applying for admission to the professional sequence of the junior year in the Elementary Education Program. The current study concerns itself with these students as they carry out the student teaching assignments of their senior year in elementary school classrooms. The setting for the study, therefore, is the Elementary Education Program at the University of Minnesota.

In the College of Education at the University of Minnesota, the program for the preparation of those who would teach in the elementary schools includes both academic offerings and courses in which students observe or lead activities with children. With the exception of a few courses in Art Education, Physical Education, and Music Education, scheduled in the freshman and sophomore year, the courses dealing directly with teaching in the elementary school are scheduled in the junior and senior years. This professional part of the curriculum is built upon an appropriate background of study in liberal arts selected for English, science, mathematics, social science, humanities, and the fine arts. Further study in the liberal arts is scheduled concurrently with the professional courses of the junior and senior year.

Two comprehensive courses, commonly known as the junior sequence, form the core of professional studies in the junior year. These courses include the study of psychological backgrounds for teaching and current curriculum practices. In both courses, students participate in nonschool activities with children on playgrounds, in settlement houses, in churches, and other organized, supervised groups designated by the College as appropriate for elementary teacher preparation. The Volunteer Service Bureau of Minneapolis and the Service Bureau of St. Paul cooperate with the College in student placement and evaluation for this work with groups. Following the two-quarter junior sequence, students enroll in a course in the philosophy of education titled, "The School and Society." This course is usually scheduled during the senior year.

In addition to the professional courses mentioned, the instruction in the teaching of reading, social studies, mathematics, children's literature, English (listening, speaking, writing), science, art, music, and physical education is included. A required course in the diagnosis and treatment of learning difficulties is scheduled concurrently with student teaching in the senior year.

The junior sequence and other teacher education courses include as a regular part of instruction visits to classrooms for the purpose of observing children at work. Students make their observations in the University Elementary School and in three Minneapolis public schools that have accepted special demonstration functions for the College of Education. These public schools are situated within a short bus ride from the college campus.

During the senior year, students majoring in Elementary Education typically reserve the morning hours of two successive quarters for off-campus student teaching in Minneapolis, St. Paul or one of ten or more suburban school systems adjacent to Minneapolis and St. Paul. Some choose to do student teaching during the first two quarters of their senior year; others select the last two quarters. Methods courses and electives remaining for the senior year may be taken concurrently with student teaching or following student teaching, except for the course in diagnosis and treatment of learning difficulties. This course is scheduled concurrently with one quarter of student teaching to allow students to study individual pupils while they are working in their student teaching assignment.

Student teachers reserve time during the late afternoon, one day a week, for an on-campus group conference with a University supervisor. The seminar type agenda for these meetings are planned by the students with the help of their supervisor.

Through the selection of appropriate courses, those who are majoring in Elementary Education may qualify to teach mentally retarded pupils or physically handicapped children. Further variation in the program is provided for students who have transferred to the College of Education with a background that includes a two-year teacher preparation, a teaching certificate, and elementary teaching experience. There is also a program for students who have baccalaureate degrees in fields other than elementary education and who, with a concentrated year in the College, may qualify for the Bachelor of Science Degree with a major in elementary education. Students enrolled in any of the variations to the regular program were not included in this study.

Students graduating from an elementary education program in the College of Education are eligible for certification to teach grades one through six in the state of Minnesota. They also may qualify to teach in grades seven and eight if these grades are included in an elementary school terminating at the eighth grade. Since students will be certified for any of the six or eight grades, the program provides that they shall have a student teaching assignment in both the primary and upper elementary grades.

Chapter II

The Design and Conduct of the Study

Introduction

To give perspective to the report of the investigation, the first chapter presented the questions to be studied, the conceptual framework, and the setting. This chapter reports the design and conduct of the study.

Although each of the four questions enumerated in Chapter 1 was of a somewhat different nature, each answer depended on the data gathered from classroom observations of student teachers. For this reason it was important to specify the sources of variability that could influence the observation data. The first division of this chapter discusses the sources of variability that were considered in collecting the data, the statistical design for the collection of observation data, and the problems of observation and measurement involved in the study. The second division, describing the conduct of the study, includes a description of the planning with the participants, the training of observers, and the methods of observation employed by observers and supervisors.

The Design of the Study

The Sources of Variability Considered

The observations of student teachers were an essential part of each of the four questions studied. Therefore, each probably source of variability affecting observational data that could be identified required a decision about how it could be incorporated into the design. Some of the sources of variability could be treated in such a way that the amount of their effect on the observations could be estimated; others could be treated only in such a way as to insure that they would not bias the results. The factors considered in the development of the design and the way in which they were treated are described under these headings: The Student Teacher, The Pupils and the Classroom, The Placement of Student Teachers in the Classroom, The Admissions Data, and The Observations by Supervisors.

The Student Teachers. One of the questions studied concerned the relationship between data collected in the admissions interview and subsequent behaviors. Therefore, the student teachers selected for this study were persons who had been subjects of the interview study. All were women in the Elementary IA curriculum of the College of Education. At the time of student teaching, all had completed at least two of their three college years in the College of Education. The fact that these were student teachers of the same sex with a common background of professional training reduced two important sources of variation.

Two other important factors concerned the student teachers. Some persons engaged in the preparation of elementary teachers have speculated that a student teacher's interest in working with a particular elementary grade level is an important factor that could influence classroom behavior.

This led to the decision that the factor of grade level preference should be incorporated in the design. The study incorporated a second factor, experience, because the usual pattern for elementary student teachers in the College of Education program is a one-half day assignment for two quarters. Student teachers available for forenoon assignments in the fall and winter quarters, 1958-59, composed the sample. The design identified the quarter of student teaching (first or second quarter) as an important possible source of variation.

The Pupils and the Classrooms. The student teaching situation received attention independent of the considerations about student teachers. In order to avoid disrupting normal student teaching activities, the study employed only those classrooms normally available for student teaching in the Twin Cities. There was no intent that classrooms should be sampled to permit generalization. As the factors taken into account are described, it is most convenient to consider pupils and classrooms as a unit.

Woven into the design of the study were two important considerations: grade level and location of schools. With respect to the first, it seemed that observed behaviors would vary depending on the elementary grade level taught. Because there is known to be a wide range of developmental differences among elementary school pupils and a diversity of materials among the elementary grades, it seemed advisable to estimate the effect that grade level might have. To reduce the variability, the upper three grades and the lower three grades formed two gross subdivisions within the elementary school.

Differences in pupils and classrooms associated with the locations of schools in different neighborhoods was the second focus for concern. Much writing in recent years suggests that socio-economic factors are closely related to pupil behaviors. Because the socio-economic factors could not be effectively incorporated into the design so that the extent of this effect could be identified, the most that could be done was to insure that this factor did not bias the observations. A socio-economic index for each school was computed on the basis of data available from the 1950 census. The index included the following information from each census tract: educational level, family income, population density, and population mobility. For each of these variables, the standard score was calculated from the city means and standard deviations. These standard scores were summed for each census tract to give its index number, then appropriately combined to give the index for a school. The use of these data are discussed in the following section.

The Placement of Student Teachers in the Classrooms. The location of elementary schools in Minneapolis and St. Paul placed some limitations on the assignments that could be given student teachers. The travel considerations for the students were of importance because all were enrolled for afternoon classes on the campus. For this reason, Minneapolis and St. Paul were divided into four geographical areas, primarily on the basis of access to schools by public transportation. The number of student teachers living in each area or having ready access to it determined the number of schools within an area to be included in the study.

The importance attached to the socio-economic similarity of these four areas is indicated in the preceding section of this report. Adjustements were made between geographical areas to insure against socio-economic bias.

When the boundaries were established, two statistical tests were made to determine whether each of these areas was generally similar to the others and to see whether the variability was similar within each. (Appendix A. See calculation of the area means and variances from school indexes). An analysis of variance test of the socio-economic index means for each area, and the test of homogeneity of the variances of these indexes showed no significant differences among the areas on these two counts. (Computations for these tests are included in Appendix A). This was the first step to insure against socio-economic bias, creating homogeneous geographical areas. The second step was also essential. Student teachers were placed so that each had an equal opportunity to be assigned to any one of the classrooms in her area. This randomization assured that chance alone, not bias, could cause the socio-economic variable to be differentially effective on classroom observations. This random assignment within geographical areas was an essential feature of the study, for it preserved the experimental nature essential to the investigation.

Admissions Data. The previous sections have identified factors that might cause a difference in the observed behaviors of student teachers. This section reviews those factors that were identified in the interview study which were used in the present investigation as predictors of observed behaviors. The predictor variables from the interview study were of two types. First were the psychological tests which the subjects completed as a part of the normal admissions procedure. These tests were the *Miller Analogies Test* (MAT), form A, the *Cooperative Reading Test*, form C2Y, the *Minnesota Teacher Attitude Inventory* (MTAI), the *Minnesota Multiphasic Personality Inventory* (MMPI), and the *Teacher Opinion Inventory*. Because these tests are either well-known standardized tests or are described in the interview study, no further discussion is presented here.

The second type of predictor variable was counselors' judgments. The interview study explored the relationship among counselors' judgments when three different bases for judgment were employed: the interview alone, test scores alone, and interview and test scores. From the analysis of the interview data and from a consideration of the procedures that counselors normally use, it seemed most logical to employ in this study the counselors' judgments based on test scores and the interview combined.

Supervisors' Judgments of Behaviors. A third type of variable essential to the purpose of this project was the supervisors' judgments based on their classroom observations of the student teachers. Because University student teaching supervisors and observers employed in the investigation made judgments in the same period, the judgments were used as a basis for comparing these methods of observation. In making their judgments, the University supervisors followed the procedure that the Office of Student Teaching normally employed in its program of supervision. There was no attempt to influence the manner in which supervisors were assigned to student teachers or the program that they followed in making their observations. The comparisons to be made required that regular procedures be maintained.

Summary. The foregoing sections contain descriptions of the factors thought to be important in the observed variables, the predictor variables and the concurrent variables in this study. These sections also identify the important sources of variation associated with the observed variables. They review the predictor variables from the interview study and describe the methods used to obtain the supervisors' concurrent judgments of classroom behavior.

The Cross-Over Design

It was the function of the experimental design to integrate into an overall plan the variables to be studied. In considering the classroom observations, it was necessary to decide which of the sources of variation would be considered experimental variables and which would remain constant. For purposes of this study the following four factors were assigned the highest priority:

1. Student teacher's grade level preference

2. Student teacher's quarter of experience

3. Classroom grade level

4. Order of placement in the two different grade level assignments.

These four factors may be considered experimental factors, for it was the purpose of the study to assess the amount of influence or effect that these variables had on the observations.

The remaining factors were considered to be controlled variables. The effect of these variables was not estimated; rather, the experiment was designed to insure that the results would not be biased by any of the four factors. They are:

1. The professional education of the student teacher

2. The sex of the student teacher

3. Travel by the student teacher

4. Socio-economic level of the classroom

The selection of student teachers for this study controlled the first two of these variables. The student teachers were women in the Elementary Education IA curriculum who had been enrolled in the College of Education for two of their three preceding college years. The division of the two cities into four geographical areas accessable to student teachers living in the area controlled the student teacher travel. Assigning schools to an area in such a way that the socio-economic characteristics of each area were equivalent controlled the socio-economic level of the classroom.

Figure 2.1 represents a diagramatic representation of the cross-over design. According to their preferences for teaching the upper three grades (prefer U) or the lower three grades(prefer L), student teachers occupied one of the two strata illustrated in Figure 2.1. Within each of the two strata, students followed one of two orders of teaching experience as determined by random assignment. One group received an upper grade assignment the first quarter and a lower grade the second. The other half of the preference group received an assignment of the reverse order. After the order of teaching assignments for each student had been determined, each person was randomly placed into a classroom of the appropriate level within the proper geographical area. Thus, the cross-over design assured that the effect of each of the four placement variables could be isolated and that those factors that were controlled would not bias the observation.

Preference	Quarter 1	Quarter 2
Prefer U	U	L
	L	U
Prefer L	U	L
	L	U

Figure 2.1 Cross-Over Design

Observation and Measurement of Variables

Of the three classes of variables considered--observed variables, concurrent variables, and predictor variables--only the latter have been specified in detail. This section discusses the observed and concurrent variables and introduces the instruments. Two different types of observation variables were required for two different purposes. These are discussed separately in the first two of the three following sections. The concurrent variables are considered in the third.

Description of Classroom Activities. One of the questions posed by this study concerned a general description of student teacher activities in the classroom and some indication of pupils' responses to student teachers. A second and related question asked whether these descriptive data were affected by the factors identified in the cross-over design.

To answer the questions, the measurement procedure needed to be objective, reliable, and interpretable. The usual approach to a description of classroom activities has been to use the ratings of supervisors or other "experts," but results from such a method are subjective, usually biased and often difficult to interpret. This part of the study required a measurement procedure that utilized the observer as one who would see and record rather than one who would interpret and judge. It was important, too, that the schedule or form upon which observations were recorded gave interpretable data concerning the activities of the classroom.

Several investigators have attacked a similar problem. Among them are H. H. Anderson and others (1945, 1946a, 1946b), F. G. Cornell and others (1952), Medley & Mitzel (1955, 1958), Withall (1949) and Flanders (1960). An investigation of these studies showed that the schedule used by Medley & Mitzel was most adaptable for the purposes given here. The Observation

Schedule and Record (OScAR) was constructed (Medley & Mitzel, 1955) by modifying the methods proposed by Cornell and Withall on the basis of try-outs of the two techniques. Modifications involved the redefinition of the scales to increase the reliability and a change in the observation method so that only one observer visited a classroom at a time. The OScAR has demonstrated sufficiently high reliability and much of its content had relevance for purposes of this study, but some modifications were made.

The Minnesota revision of the OScAR (OScAR III) and the manual for its use are in Appendix A. The content of the revision is similar to the original in several important ways. Nearly all of the items in the "Activities" section of the OScAR 2a, Medley & Mitzel's revision of their original, were retained unchanged. The "Signs" section was omitted and most of the items in that section were incorporated into the "Activities" section. The "Materials" section was changed by separating teacher and pupil materials into two sections. The "Subject" section remained substantially the same.

The OScAR III incorporated two major changes: the first substitued an entirely new section that provided a method of recording ways in which groups were formed, their size, whether or not the teacher was a member of the group, and whether the individuals worked individually or cooperatively; the second omitted a section providing for the classification of teacher verbal statements. Because the classification of verbal behaviors was of a special concern to the major questions of the study, this problem was considered somewhat separately from that of describing the activities of the classroom. The following section gives this matter separate treatment.

The method of recording observations for the OScAR III was exactly the same as used by Medley & Mitzel. Observers checked whether each behavior occurred in each five minute observation period. In a sense, each observation category or item is analogous to a test item which is answered "yes" or "no" for each five minute observation period.

Chapter 3 presents the scales resulting from the grouping of the OScAR III items and the resulting scores for individuals in summary form. Succeeding chapters present an analysis of the effect of other variables on these scales.

Classifying Classroom Verbal Behaviors. The previous chapter presented the behavioral characteristics which were assumed to be evidence of the constructs of need Dominance (n Dom) and need Integration (n Intg). It also discussed the relationship of each of these constructs to classroom behaviors. In this section the purpose is to describe the method of observing classroom verbal behaviors and the treatment of these observations to obtain a measurement of the variables.

Table 2.2 presents the categories which this study used in classifying classroom verbal behavior. Flanders defined the categories and developed them as a part of the Minnesota System of Interaction Analysis (Flanders, 1960). Although the published version reflects some minor revisions in the categories as used in this study, these changes do not affect the interpretation of the results of the study. The table indicates that there are four categories of indirect influence and four categories of direct influence

Table 2.1

Categories for the Classification of Teacher Verbal Behaviors
(From the Minnesota System of Interaction Analysis, 1958)

		Category Number	Category Description
TEACHER TALK	INDIRECT INFLUENCE	1	Reacts to student feeling: responds to the emotional tone of a student's behavior in a positive, non-threatening way, e.g., "you are worried about this one, aren't you?"; "are you wondering how you can ever finish the assignment?"
		2	Praises or encourages: complimenting, commending or praising the work of a student; providing tension release for the class to maintain or improve the classroom atmosphere; encourages to keep talking.
		3	Reacts to student idea: responds to the content of pupil's behavior, rather than the affect, in a positive and nonthreatening manner.
		4	Asks a Question: asks about procedures or about the content of the material being studied with clear expectation of an immediate answer.
		5	Neutral statements: administrative or routine statements that are not concerned with present learning activites.
	DIRECT INFLUENCE	6	Gives information or opinion: lecturing, stating opinions or facts, asking categorical questions with no expectation of immediate response.
		7	Gives directions: requests or commands where normal student compliance is expected; confirm or approve student compliance.
		8	Criticizes: remarks intended to change the present undesirable behavior or attitude of a student to a more desirable one, including use of sarcasm, scorn, ridicule, laughter.
		9	Self-justification: remarks by teacher explaining teacher behavior, justifications of teacher control, excessive self reference.
STUDENT TALK		10	Student talk-response: talk by students in response to teacher. Teacher initiates the contact or solicits student statements.
		11	Student talk-initiation: talk by students which they initiate. Student may by "called on" if teacher is giving student his turn.
		12	Dilema or confusion: pauses, short periods of silence or periods of confusion when communication cannot be understood.

*No scale value is implied by the category numbers. They are used only to identify the categories.

that describe the teacher's communication to the class concerning the primary learning tasks. One additional category is used for teacher talk that relates to administrative or routine procedures which are not directly related to the primary learning tasks of the class. Pupil talk is characterized as being either self initiated or as teacher initiated, a specific response to the teacher. The purpose of these categories in the Minnesota System is to include all possible communication events that might happen in the classroom. An observer records the events of the classroom in sequence, making an observation twenty to twenty-five times a minute. In the Minnesota System the observer uses the categories which best describe the communications events in the approximately three second interval between observations. These data are then recorded in sequence to permit classification of successive pairs of events. In this way the Minnesota System gives a picture of the patterns of categories that a teacher uses from time to time and from situation to situation. The comparisons made according to this system are primarily intra-teacher comparisons.

In this study, the method of sampling the verbal behaviors and summarizing them were somewhat different than used in the Minnesota System. Using the categories specified in Table 2.1, observers recorded each category a teacher used rather than the one which summarized each three second period. Observers recorded each change of communications events regardless of the interval of time between them. When the same type of event persisted for more than the three second period, the observer continued to record the use of the appropriate category each three seconds until the category changed. Such an adaptation of the Minnesota System gives a picture of the frequency or relative use of the categories in the observation period, rather than giving an indication of the relative amount of time during the observation period which was devoted to each category.

In this study the method of summarizing data from an observation period did not describe, as the Minnesota System does, the chain of communication events. Here the purpose was to represent the degree to which n Dom and n Intg were evident in the teachers' behaviors by the extent to which the direct and indirect categories of influence were used.

The numerical representation of these variables was the sum of the number of times each of the twelve categories of interest occurred within each five minute observation period. This gave the total number of communication events which occurred during the period. This number tended to be 100 with slight variations between different observers for the same period and for the same observer during different periods. To adjust for these minor variations, which occurred primarily because of slight differences in estimating the three second intervals used when categorizing successive communications events of the same type over a period of time, the sum for the five minute period was adjusted to a base of 100. The numbers representing the use of each of the categories were similarly adjusted to this base, which produced a quantity for each category representing the frequency of its occurrence per hundred tallies during that period. The variables of concern, n Dom and n Intg, were then represented for that five minute period as the sum of the frequency per hundred for the four relevant categories. Thus the sum of the observations in categories six through nine represented the variable n Dom, and the sum of categories one through four represented the variable n Intg. In each visit to a teacher the observers recorded the verbal behaviors over three separate five minute observation periods. The score representing the visit of the observer was thus the sum of the scores obtained for the separate five minute periods.

Although one can give strong support to the argument that the occurrence of the categories ought to be weighted differentially according to their psychological contribution to the concept being measured, there seemed to be no defensible rationale for determining an a priori weighting scheme. Weightings were made, however, using the differential contribution of each category to the total score by means of reciprocal averages. Chapter 3 presents a more detailed description of this procedure.

Judgments of Behavior by Supervisors. The third question for study cited in the first chapter concerned the relationship between the ratings made by experienced student teaching supervisors using their methods of observation and the data obtained by the research observers using the methods of observation designed for this study. The question was stated in alternative forms that differed according to the measurements used as the basis for comparison. The first form referred to the concepts of n Dom and n Intg as the bases for comparison and the second form concerned some objectives of the student teaching experience. In either case, the comparisons involve research observers using the methods of observation previously described and experienced supervisors using an approach to observation consistent with their purposes. Whatever the comparisons, one cannot extricate the separate effects of the experience and background of the rater from the methods of observation he employed, for these two factors are inseparable in this study. Because the data from only one academic quarter were necessary to the comparisons, comparative data were gathered for the second quarter only.

The comparison of the different observers and methods on the basis of the constructs of n Dom and n Intg employed the following method. For the research observers the variables derived from the categorizing of verbal behaviors as described in the previous section were used. For the supervising teachers, the variables were determined by their ratings on the form that appears in Appendix A. After the supervisors had completed their observations as they would normally conduct them, they completed a rating on the n Dom and n Intg scales. Each supervisor made his rating on the five point scale and was instructed that the descriptions of n Dom and n Intg that appear in the first chapter represented the highest values on the scales. Each supervisor completed at least three observations on the same student over the course of the quarter. Comparisons were thus based on the composite scores derived from categorizing verbal bahaviors, and the composite judgments indicated by the scale values assigned.

The second form in which question three appears required the construction of an instrument having relevance for the objectives of student teaching. It was necessary to limit these objectives to those aspects of student teaching which could be observed while the student teacher was in charge of the class. Certainly there were other student teaching activities, for example, those which required written preparations or cooperation with supervising teachers and University supervisors, that those who observe only the classroom teaching behaviors could not evaluate. Since question three dealt with a comparison that involved only the classroom teaching behaviors, other objectives of student teaching were excluded from consideration.

To assess the objectives that could be observed in classroom behavior, the supervisors and the observers completed a twenty-one item scale. The scale contained two sections. The first, headed "Creativity and Resourcefulness in Teaching," included thirteen items, and the second, headed "Interpersonal Relations," contained eight items. The scale was adapted from a longer one developed by the Coordinator of Elementary Education and her staff of student teaching supervisors for use in the program of the College of Education, and was, therefore, particularly appropriate for the students in this study. Some additions were made, particularly in sections relating to creativity, in order to include items tapping all the dimensions of creativity suggested by Lowenfeld (1958). In addition to specific items, the scale included a request for an overall estimate of success in student teaching and an estimate of probably success in teaching, each to be rated on a five point scale. A copy of this instrument, "Ratings of Classroom Behavior," is included in Appendix A. The study includes separate computations for the two sections and for the summary estimates. Both the observers and the supervisors completed the ratings after finishing their normal periods of observation.

The Conduct of the Study

Planning with the Participants

Public Schools. As in any such projected study involving the public schools, consultation started with the appropriate school officials. After discussion of the tentative plan, each official gave permission for the research project directors to request cooperation from principals and teachers. The directors followed the procedures suggested by school officials as they proposed the research and requested permission to make classroom observations from those who were available to supervise student teachers.

In retrospect, an orientation meeting was the most satisfactory means used to enlist the cooperation of supervising teachers. In these meetings, project staff members held informal discussions with teachers before the student teachers were placed. Such meetings gave teachers an opportunity to hear a firsthand account of the project and to have their questions answered. When the investigators relied on other methods of communication, teachers expressed much concern. In one case a student teacher arrived in the classroom to participate in this research project before the teacher was acquainted with it. Fortunately such incidents were few, and when questions arose teachers were understanding and helpful. Even though meetings of cooperating teachers were difficult to arrange, the results justified the extra effort. In this project those teachers who did not have an opportunity to participate in the discussion of the project before the observations began were invited to such a meeting early in the fall quarter. Research staff visited individually with teachers who could not attend group sessions.

Student Teachers. After the directors were assured of the participation of the schools and the approval of research funds, they called a meeting of those student teachers who were available for placement in the Twin Cities for the fall and winter quarters, 1958-59. The purpose of the meeting was to explain to them the purposes of the research project. University supervisors of the student teachers also attended this meeting. The procedures for scheduling research observations were explained to the group. In general, the group showed an interest in the project and a willingness to cooperate.

Training the Observers

The primary consideration in the selection of research observers was their sensitivity to interpersonal relations. The persons selected had taken course work in psychology and the behavioral sciences and their background evidenced some experience in dealing with individuals and groups. None of those selected had been prepared to teach or had taught in elementary schools, but some had taught in the secondary schools or in college. Others were experienced counselors and personnel workers. Because neither of the observation methods required a background in elementary school teaching, it did not seem essential that observers have this type of preparation, and for the sake of generalizability of the observation procedures, it seemed best that such experience not be required.

During the four weeks prior to the first observations, the seven observers, who were hired on a half-time basis, spent a total of 60-65 hours in training in the techniques of observation. They spent the greater part of this time learning to categorize verbal behaviors as outlined by the Minnesota System of Interaction Analysis. Although the system of categorization is somewhat complicated and takes concentrated training for efficient and reliable observing, the length of time set aside for training was more than ample. The time alloted permitted sufficient opportunity to use the technique in a variety of settings, and insured that observers were indeed "like-minded" in their use of the system.

When the training in the categorizing of verbal behaviors began, Professor Ned A. Flanders, who developed this system, spent several mornings with the group acquainting them with the categories, their rationale, and the method of recording observations. In the initial days of training the observers used tape recordings of classroom sessions. Following this, the observers moved into University Elementary School and nearby public schools to practice in a "live" situation. Following each period of observation, the observers discussed the problems that arose in applying the categories to the communications events. The following chapter discusses the method used to calculate the reliability of observers and presents the results of the reliability calculations at various points in the training of the observers.

Training in the use of the OScAR III was not introduced until the latter part of the initial training period. Because this observational method depended primarily upon a familiarity with the categories, less training was required to become reliable in its use. The detailed manual for the OScAR III made the use of the method clear and, in effect, shortened the time necessary for training the observers in its use. In addition, the reliability of this method had been established by Medley and Mitzel.

The last phase of the training integrated the use of the OScAR III and the Minnesota System, as in the conduct of the study. Beginning with the OScAR III, the observers used each method for a five minute period, alternating them for a total of thirty minutes. Thus there were three periods in which they used the OScAR III and three periods in which they used the verbal categories.

After the observers began their observations of the student teachers, they made periodic reliability checks by going as a group to some nearby elementary classroom, as they did during the training session. They made such reliability checks weekly at first, then every other week. During the period between the two quarters, the observers engaged in training sessions to maintain the level of skill that they had developed.

Assignment of Student Teachers

In the original plan for this study, the directors estimated that sixty students from the 167 students who participated in the interview study would be available for this study. After deciding which two quarters were to be used for observatios and listing the classrooms available, they found only forty students were available for assignment according to the experimental design.

The student teachers were assigned in the following manner. A project staff member first plotted on a map of the Twin Cities area the residences of those students who were thought to be available for this study to determine their relationship to the schools having classrooms available for the assignment of student teachers. After checking the routes of the public transportation system, the geographical areas were tentatively outlined so that the student teachers would have reasonable access to schools within a given area. Using the census data previously described, the area boundaries were then adjusted so that the socio-economic characteristics of each area were as nearly alike as possible, and so that each area contained an equal number of upper grade and lower grade classrooms. When the classrooms were thus determined, they were randomly paired by level and order, to determine the two-quarter classroom assignments patterned according to the design. Finally, the students were assigned randomly within each geographical area after they had been stratified according to their grade preference.

As might be expected, it was necessary to make some adjustments. Two students were unavailable for assignment, one because her low grade point average disqualified her for a student teaching assignment, and the other because she chose to postpone student teaching. Thirty-eight students then made up the sample. After the thirty-eight students were assigned, two found that they had class conflicts that could not be reasonably solved. Therefore, two student teachers who were originally placed in suburban schools, and not included in the randomization, though eligible, were substituted. After the study began, two supervising teachers asked to be relieved of students teachers, one because of illness, and one because of the pressures of the extra assignment. Substitutions were made for these teachers without disturbing the pattern required by the experimental design. Of the thirty-eight student teachers, two were dropped from the sample before the end of the study. One had inadvertantly been assigned incorrectly and the second had to be moved from an undesirable assignment for her own welfare. The investigators were careful to protect the requirements of the design, so that the only case which was not treated in keeping with the requirements was dropped from the study.

Conducting the Observations

Each of the five observers visited each student teacher once during each quarter. Thus, there were ten different observation reports for each student teacher.

In the fall quarter there were thirty-five days when it was possible to conduct observations. The thirty-eight student teachers were assigned to twenty-nine different schools in Minneapolis and St. Paul. Seven research observers (two of whom were employed in the project the first quarter only) had to be scheduled to visit each student teacher. The visits to each teacher had to be spaced over the quarter. These plans allowed for a two week period at the beginning of the quarter for the student teacher to get acquainted with the class before observers visited her.

The fall quarter schedule of observation allowed thirty-two days to complete the data collection and provided three additional days for any observations that had to be made up. The plan called for an average of eight and three-tenths observations per day with three to four days between two observations in the same classroom. Actually, the time between observations ranged from zero to fourteen days. In one or two instances the illness of a student teacher or the supervising teacher at the time when observations were planned made it necessary to complete two observations in one morning period later in the quarter. In no case, however, were two observations made simultaneously.

Originally the visits of the research observer were planned week by week on the basis of the student teacher's plans for teaching, which had been made with the supervising teachers. As the investigators sought to improve methods of scheduling, the supervising teachers suggested that the schedule of visits be established by the research project coordinators. Teachers and student teachers found it easier to plan their activities around the scheduled visits than to predetermine when the student teacher would be teaching and to communicate this to the project coordinator. As was previously indicated, at the end of the fall quarter it was necessary to drop two student teachers from the study because of scheduling difficulties. It was also necessary to reduce the staff of research observers from seven to five.

During the winter quarter the thirty-six subjects were assigned to twenty-seven schools in Minneapolis and St. Paul. Allowing for vacation days and an adjustment of time between the first day in the new assignment and the first observation, there were forty days available for observation during the winter quarter. Data were scheduled to be collected in thirty-seven days. The time called for approximately five observations per day, each student teacher to be visitied once a week. By actual count there was an average of 4.8 observations per day with an average of 6.3 school days between observations. Time between observations ranged from zero to twelve teaching days.

After the experience of administering this experiment, the investigators reached two somewhat different conclusions. First, there may be a very good reason why this study is probably the first of its kind, that is, one that employed an experimental approach to the placement of student teachers in classrooms: it's a big job! The organizational and administrative tasks were complex and time consuming. Second, the tasks of such a project are not insurmountable; research of this type is feasible as well as important.

Chapter III

The Observed Behaviors of Student Teachers

In the previous chapter the observational methods were introduced and the problems in their use were discussed. In this chapter the data collected from the observations are presented. The methods of scoring and scaling are given, as is a general description of the data. The data in the first half of this chapter give answers to two of the questions posed in the initial chapter regarding the activities of student teachers and pupils responses to them.

Data from the OScAR III

Scoring and Scaling

Each OScAR III sheet recorded the observations made in each of the three alternate five minute periods. Each observation was represented by the total of the three five-minute time periods. A preliminary analysis indicated that some of the individual items were rarely or infrequently used, so some arbitrary groupings were made to facilitate further data processing. Figure 1 indicates the item numbers given to each observational category or to a group of categories. The item data were coded and punched into IBM cards. For items that consisted of only one observational category the values representing the actual frequency of occurrences were 0, 1, 2, or 3. For items that were composed of a group of categories a 0, 1, 2 . . ., 7 scale was used. For these groups of categories the item values 1 to 6 indicated the actual total frequency of occurrence. Those that totaled seven or more were coded seven. Appendix B contains the table which reports the frequency of occurrence of each of the items.

The grouping of individual items into scales was done largely on a a priori basis, although the organization of items in the OScAR and the factor analysis of Medley and Mitzel (1958) did influence the decisions which designated the eight scales. Table 3.1 presents the scales and the items comprising each scale. Figure 3.1 presents the form of the OScAR III, and each of the items is keyed to the scales presented in Table 3.1.

The Univac Scientific Computer (ERA 1103) was used to obtain each scale score for every observation for all student teachers. The simplified test analysis program developed by Baker (1959) computed the score for each of the ten observations, and calculated the mean and standard deviation for each scale. So that individual scores on various scales would be numerically comparable, a linear transformation was performed which set the mean scale value at 50 and the standard deviation at 5. The actual mean and standard deviation values for each of the scales is given in Table 3.1. Inasmuch as the reliabilities associated with this method of observation were established by Medley and Mitzel (1958) as ranging from .60 to .89 no reliability determinations were made in this study.

Figure 3.1 - Items and Scales for OScAR III

	Item No.	Scales	
I.			
A. 1. t qu, p ans	1	TXA	CVE
2. t ans p qu	2		
3. t intrrpts p	3		
4. t ignrs p qu	4		
5. t wrns, thrtns p	5		CEC
6. t crits prrcm	6		
7. t prs pr gd wk	7		
8. t prs p gd bhv	8		
9. t wks w ind p	9		
B. 1. t lctrs	10	TCA	CVE
2. t rds, tls stry	11		
3. t tks t cl	12		
4. t ills at bd	13		
5. t ills wth mp, cht	14		
6. t dmnstrs	15		
7. t shws flm, sld, pls rcd	16		
8. t pss bk, ppr, mlk	17		
9. t lds sng, ex, gm	18		
C. 1. t wks at dsk	19		
2. t clns, dcts rm			
3. t wrts on, ds o bd			
4. t lvs rm			
5. t tlks wth vstr			
II.			
K. 1. p whsprs	20	PDB	CEC
2. p lghs	21		
3. p intrrpts t	22		
4. p igns q of t	23		
5. p shws hst t t	24		
6. p shws af t t	25		
L. 1. p tks t gp	26	PCA	CVE
2. p rcts	27		
3. p rpts, gve prpd tk	28		
4. p rds ald	29		
5. p dmnstrs, illus	30		
6. p gvs skt, ply	31		
7. p sngs, pl instr	32		
8. p pls gm	33		
9. p lds cls	34		
10. p lds sng, ex, gm	35		
11. p pss bks, ppr, mlk	36		
M. 1. p rds, stdys at st	37		CVE
2. p wrts, maps at st			
3. p pnts, cts, drws	38		
4. p wks at bd			
5. p dcts rm, bd	39		
6. p clns rm, bd			
7. p rsts, hs snk			
8. p lvs rm			
N. 1. p scffls, fghts wth p	40		CEC
2. p shws hst t p	41		
3. p shws aff t p	42		

*

**

	Item No.	Scales	
III.			
V. 1. t chs vlntrs	43	TGA	
2. t chs			
3. p chs vlntrs	44		
4. p chs			
5. p t gp of chs			
W. ind tsk w/t			
1. over ½ cl wt	45		
2. 4-½ cl wt			
3. 2-3 wt			
X. gp tsk w/t			
1. ovr ½ cl wt	46		
2. 4-½ cl wt			
3. 2-3 wt			
Y. ind tsk w/o t			
1. over ½ cl w/o t	47		
2. 4-½ cl w/o t			
3. 2-3 w/o t			
Z. gp tsk w/o t			
1. ovr ½ cl w/o t	48		
2. 4-½ cl w/o t			
3. 2-3 w/o t			
IV.			
Teacher Materials	49	CUM	CVE
1. blbd			
2. mp, cht, pctr			
3. sld, fm, etc			
4. audio aid			
5. 3 dim object			
6. txt, wkbk	50		
7. supl rdg mtr			
8. wrtng			
9. hcft, art	49		
10. spec tchg aid*			
11. no mtls usd			
V.			
Pupil Materials			
1. blbd	51	CUM	CVE
2. mp, cht, pctr			
3. sld, fm, etc			
4. audio aid			
5. 3 dim object			
6. txt, wkbk	52		
7. supl rdg mtr			
8. wrtng			
9. hcft, art	51		
10. spc lrng aid**			
11. n mtls usd			
VI.			
1. rdg	53		CVE
2. math	54		
3. lang arts	55		CVE
4. soc st	56		
5. science	57		
6. recreation	58		
7. art, craft	59		
8. music	60		
9. test			
VII.			
1. sprvsng tchr prsnt			
2. sprvsng tchr cmmnts			

Table 3.1

Item Numbers, Means and Standard Deviations for Eight Scales Derived from the OScAR III

Pupil Disruptive Behavior	20, 21, 22, 23, 24	5	49.8	5.0
Pupil Class Activities	26, 27, 28, 29, 30, 31 32, 33, 34, 35, 36	11	50.2	4.6
Teacher Class Activities	10, 11, 12, 13, 14 15, 16, 17, 18	9	51.0	4.1
Teacher Grouping Activities	43, 44, 45, 46, 47, 48	6	49.8	3.0
Teacher Individual Activities	+1, 2, 7, 8, 9 -3, 4, 5, 6	9	53.4	3.5
Classroom, Variety of Materials	49, 50, 51, 52	4	51.1	3.8
Classroom Emotional Climate	+7, 8, 25, 42 -5, 6, 20, 21, 22, 23 24, 40, 41	13	50.0	5.2
Classroom Verbal Emphasis	1, 2, 10, 11, 12, 26, 27, 28, 29, 37, 50, 52, 53, 55	14	50.0	4.3

Student Teacher Activities

In this section of the report the purpose is to present a general description of the subject areas in which the student teachers taught, the materials and the procedures they used. Headings for this discussion are six of the eight scales derived from OScAR items. Since Classroom Emotional Climate (CEC) and Classroom Verbal Emphasis (CVE) are composed largely of items which overlap other scales they will not be discussed here. It should be noted that the observation methods employed in this study do not permit statements concerning the amount of time spent in the various activities. The percentages used represent the proportion of visits in which the activity or behavior was observed. Such figures can be used only as gross indications of the extent to which the activity was actually present. The relationships among the various descriptive scales is presented in the final section.

Subject Areas Taught

Because all of the student teachers in the study were in the classrooms during the half day before noon only, their activities would be expected to be similar in nature. To be able to make a more precise statement about these activities, the subject or subjects taught were recorded during each observation period.

As might be expected, reading was the subject most often being taught during observation periods. Forty-five percent of the observations recorded the teaching of reading. Social studies was observed taught in slightly more than one-third of the periods. Only two other subjects, language arts and arithmetic, occurred more than one period in five. Art and music teaching were observed in ten percent of the visits, science in only seven percent.

Variety of Materials Used

The check list used for recording the observations contained two sections where the materials used were noted, one section for pupil materials and one for teacher materials. For purposes of analysis, the items in each section were grouped into two categories. One group contained materials, such as texts, workbooks, and writing materials, which were called "verbal materials." The other category included such audio-visual materials as blackboard, maps, films, records, and special objects or learning aids, and were labeled "audio-visual materials." The scale was labeled Classroom, Variety of Materials (CVM), and was composed of the total of teacher and pupil materials used during the observation.

Student teachers were observed using verbal materials in three out of four of the observations. Even more frequently, 80% of the visits, audio-visual materials were seen in use. Pupils were using verbal materials in nearly all, 86%, of the observations, but were recorded employing audio-visual materials in but slightly over half, 56%, of the periods.

Teacher Class Activities. The OScAR III contained a series of items that were checked when the student teachers took an active lead in directing the class. In these activities the pupils were more or less passive participants. Examples of the activities are these: Teacher lectures, tells stody, reads to class, illustrates at board, demonstrates, leads singing or games.

The most frequent item observed was the student teacher talking to the class. Observers checked this item when the teacher discussed nonacademic behaviors with the class. Such topics as classroom or school conduct, school patrol topics or other such school activities were included in this item. In nearly all of the visits (96%) "teacher talks to the class" was observed. A formal presentation of subject material, activity resembling a "lecture," was observed in about half (52%) of the visits. Illustrating at the blackboard was noted in half of the observation periods, but the use of maps and charts was seen, on the average, only once in five periods.

Teacher Grouping Activities. One section of the observation schedule was designed to record the formation of groups by the student teachers. Certainly grouping students is one of the important ways to individualize the learning process. Although the use of this technique by student teachers may be less effective than its use by the regular classroom teacher, the only question asked in this study concerned the extent to which this teaching method was used by student teachers in the sample. In slightly more than one in ten observation periods (13%) was a group formed by the student teacher. Almost always student teachers decided which pupils formed the groups. Less than one group in ten formed by the student teachers came together on the basis of pupils' choosing the members of the group. In almost all of the groups that were formed each pupil was working on an individual task of his own that did not require reference to other group members.

Teacher-Individual Activities. To try to describe the activities of the teacher which were directed toward individual pupils, one section of the OScAR III contained items describing these kinds of activities. These items were ones that would be helpful in indicating what student teachers do to encourage pupils individually in their learning efforts. The more of these activities present, the more the student teacher could be described as inquiring, helpful, patient, and supportive of the pupils.

The most common method of influencing individuals was by direct questions to the pupil, questions requiring a response. This type of questioning was observed in almost all of the periods. Answering individual pupil's questions, rather than ignoring them, was seen in eight out of ten visits.

Pupil-Class Activities. In one section of the schedule observers recorded the events in the classroom in which individual pupils were the focus of class attention. Pupils may have been reciting, reporting to the class, telling stories, or leading the class in a number of activities. The acitvities are similar in nature to those in which the student teacher leads the class (TCA) except that the leadership comes from the pupils rather than from the student teacher.

The most common of these activities was pupil recitation, occurring in 77% of the observation periods. Observers were instructed to check this category when pupils were asked to recite materials that were to have been learned, for example, "the most important minerals in Minnesota." In over 60% of the periods, records showed pupils talking to the group; for example, giving extemporaneous descriptions of a recent event. In over half of the periods, pupils read aloud. Pupils gave demonstrations in one out of four periods but led the class in an activity, game, or song, only once in twenty periods.

Pupil Response

No doubt there are many ways that could be chosen to characterize pupil response in the classroom. For this study the authors were interested in observing how cooperative pupils were as student teachers went about their tasks of learning the art of teaching. That pupils "test" their student teachers is well known, and if these student teachers have listened to the tales of the student teachers of other years they, no doubt, expected to be "tested." Two purposes of the study were to determine to what extent pupils were disruptive, and to find the factors that were associated with disruptive pupil behavior.

The observation schedule which was used in this study permitted the observer to record a number of "pupil disruptive behaviors." The instructions given to the observers were that these behaviors must clearly indicate that the pupils were inattentive when attention was expected, reluctant to cooperate, or resistant to student teachers' efforts. Whispering, laughing, ignoring requests, interrupting, and hostile behavior toward the student teacher were recorded. Some of these behaviors can certainly be regarded as normal in the sense that they are characteristic of certain age groups in the elementary school. Recognizing this, it was the observer's purpose to distinguish when these behaviors were normal and appropriate and when they were disruptive.

Of these disruptive behaviors, unauthorized whispering was the most common, occurring in 54% of the observation periods. In nearly one-half of the periods observers recorded pupil interruptions of the teacher and in one-third they observed disruptive laughter. Much less frequently, once in ten visits, observers recorded pupils ignoring teachers' requests or showing outright hostility.

Relation of Student Teacher Activities to Pupil Response

Since one of the five descriptive scales constructed from the OScAR III concerned pupil response (PDB) and four concerned activities initiated by the student teacher, it was of interest to see how these scales correlated. Table 3.2 presents these correlations. The table also presents four other correlations of interest which will be described below.

Table 3.2

Intercorrelations of OScAR III Scales Describing Classroom Activities and Pupil Response

	(2) TCA	(3) TGA	(4) PCA	(5) CVM	(7) PDB	(8) CEC	(9) CVE
(1) TIA	.00	.01	.22	.31	-.63**		
(2) TCA		.11	.25	.22	-.22		
(3) TGA			.20	.27	-.20		
(4) PCA				.10	-.30		
(5) CVM					-.16		
(6) CAI					-.49**	.56**	.36*
(8) CEC							.06

*significant at .05 level
**significant at .01 level

Five of the scales, TIA, TCA, TGA, PCA, CVM, were described earlier in this chapter. Table 3.2 shows that none of the intercorrelations among these scales was significantly larger than could be expected by chance. The low intercorrelations of each of these scales indicated that each measured a different kind of classroom activity. For further study of the interrelations of the OScAR III variables, the five scales mentioned were combined into a single Classroom Activities Index (CAI). Since each of the individual scale scores reflected more the variety of activities and only roughly the amount, CAI may be considered as an indication of the total variety of classroom activities instituted by the teacher.

When each of the six scales, the five individual scales and the composite scale, were correlated with Pupil Disruptive Behavior (PDB), one important relationship was found. All the correlations were negative, two significantly larger than would be expected by chance. Even though four of the correlations did not reach normal significance levels, it is noteworthy that all of the activities were correlated negatively with disruptive behaviors. In general, then, the more frequently a variety of activities was observed, the less disruption by pupils was seen.

The relationship between TIA and PDB, -.63, was the strongest relationship found. In those classrooms where student teachers were more frequently observed to be interested, helpful, and supportive, the disruptive behaviors were less. The relationship between CAI and PDB indicates that the combination of single scales is significantly related to disruptions even though four of the five individual scales comprising CAI do not indicate a relationship when considered singly.

Two scales, Classroom Emotional Climate (CEC) and Classroom Verbal Emphasis (CVE) were suggested by Medley and Mitzel's factor analysis of classroom observation data. These scales contain a large percentage of item overlap with the five scales discussed above. For example, four of the twelve CEC items are found in the CAI scale, and all but three of the CVE scale items are included in CAI. It is not surprising, then, that CAI should correlate significantly with CEC and CVE. There are no overlapping items between the CEC and CVE scales. These two independent scales were found to be unrelated, r = .06. It is interesting to note that the emotional climate of the classrooms was independent of the extent to which a verbal emphasis was found in the classroom activities.

Data from the Classification of Verbal Behaviors

Chapter I presented discussion of the concepts of need Dominance and need Integration and their relevance for this study. Chapter II described the Minnesota System of Interaction Analysis as it was used. The purpose here is to present the reliability determinations for the n Dom and n Intg variables and the comparison of weighted and unweighted scoring methods.

Observer Reliability

The experience of Medley and Mitzel suggested that the most efficient method of classroom observation was to send one observer to a classroom at a time. As a result the deisgn of this study could not incorporate a satisfactory method for determining the reliability directly from the observations. The reliability of the OScAR technique was well established, but some method was needed for giving an estimate of the reliability of classifying teacher verbal behavior. Because of the different use of scores in this study from that normally used in the Minnesota System of Interaction Analysis, and because of the method of assigning observers, a means of establishing reliability apart from the experiment was employed. During the training sessions prior to the observations made for the study, and at regular intervals subsequently, all observers were sent to observe the same teacher or teachers. From these observational data, estimates of observer reliability were made.

In general, determination of the reliability estimates were made from an analysis of variance model which specified the sources of variation which could affect the scores which resulted from the classification of observed behaviors. Each source of variation was designated as contributing either to the true differences among observed scores or to error. The reliability estimates for an average judge were calculated by finding the proportion of the variance of observed scores that could be attributed to true differences in scores.

More specifically Table 3.3 presents the analysis of variance model for determining the reliability of an average judge when several judges observe one teacher. The table also specifies the expected mean square values for each source of variation and designates whether it contributes to true score or error. The model assumes that (1) the observed scores result from a linear combination of the factors specified, (2) observers and observational periods are random variates, and (3) the categories used in classifying verbal behaviors are fixed variates. It is further assumed that the successive judgments made by the same observer are correlated. The partition of the sum of squares follows that which is characteristic of the split plot designs. Table 3.4 presents the partition of the sum of squares and the expected mean square estimates when the team of observers repeats its observation procedures in the classrooms of several different teachers.

The reliability coefficients computed from data gathered during thirteen observations, three training sessions before the beginning of the experiment and ten during the experiment or between fall and winter quarter, are presented in Table 3.5. The reliabilities of an average judge based on scores of n Intg ranged from .65 to .86 with a mean of .77. For scores on n Dom the values ranged from .65 to .93 with a mean of .81. The table also shows the reliability coefficients for the pooled ratings of five judges, adjusted by use of the Spearman and Brown prophesy formula. For n Intg the r's range from .90 to .97, and for n Dom the values range from .90 to .98.

Since observed variables used in the analysis of the experimental data were composite scores representing the pooled observations of five judges, it may be inferred that the score values computed for each subject had sufficient reliability to be stable estimates of the variable observed.

Table 3.3

The Partition of the Sum of Squares for Determining the Reliability of an Average Judge When Several Judges Observe One Teacher at Different Time Periods

Designation	Source of Variation	Degrees of Freedom	Expected Mean Square
Error	Observers(o)	I-1	$S^2(ocp) + J\ S^2(op) + JK\ S^2\ (o)$
True	*Categories(c)	J-1	$s^2(ocp) + I\ S^2(cp) + K\ S^2(oc) + IK\ S^2(c)$
Error	Observer x Category	(I-1)(J-1)	$S^2(ocp) + K\ S^2(oc)$
True	Periods(p)	K-1	$S^2(ocp) + J\ S^2(op) + IJ\ S^2(p)$
Error	Observer x Periods	(I-1)(K-1)	$S^2(ocp) + J\ S^2(op)$
True	Category x Periods	(J-1)(K-1)	$S^2(ocp) + I\ S^2(op)$
Error	Observer x Category x Periods	(I-1)(J-1)(K-1)	$S^2(ocp)$

*Fixed variate

Table 3.5

The Reliabilities for the Average Judge and For a Composite of Five Judges Estimated from a Series of Thirteen Observations Over a Two Quarter Period for n Intg and n Dom Variables

Observation	n Intg		n Dom	
	Average	Composite	Average	Composite
1	.75	.94	.88	.97
2	.73	.93	.85	.97
3	.76	.94	.70	.92
4	.65	.90	.88	.97
5	.81	.96	.81	.96
6	.86	.97	.93	.99
7	.86	.97	.84	.96
8	.77	.94	.65	.90
9	.85	.97	.84	.96
10	.80	.95	.72	.93
11	.68	.91	.89	.98
12	.73	.93	.74	.93
13	.72	.93	.81	.96

Table 3.4

The Partition of the Sum of Squares for Determining the Reliability of an Average Judge When Several Judges Observe One Teacher at Different Time Periods Replicated for Several Teachers

Designation	Source of Variation	Degrees of Freedom	Expected Mean Square
Error	Observer(o)	I-1	$S^2(poc/t) + J\ S^2(po/t) + JK\ S^2(ot) + JKL\ S^2(o)$
True	Teachers(t)	L-1	$S^2(poc/t) + J\ S^2(po/t) + IJ\ S^2(p/t) + JK\ S^2(ot) + IJK\ S^2(t)$
Error	Observer x Teacher	(I-1)(L-1)	$S^2(poc/t) + J\ S^2(po/t) + JK\ S^2(ot)$
True	*Categories	J-1	$S^2(poc/t) + I\ S^2(pc/t) + IK\ S^2(ct) + K\ S^2(cot) + KL\ S^2(co) + IKL\ S^2(c)$
Error	Category x Observer	(I-1)(J-1)	$S^2(poc/t\_ + K\ S^2(cot) + KL\ S^2(cot)$
True	Category x Teacher	(J-1)(L-1)	$S^2(poc/t) + K\ S^2(cot) + IK\ S^2(ct)$
Error	Category x Teacher x Observer	(J-1)(L-1)(I-1)	$S^2(poc/t) + K\ S^2(cot)$
True	Periods Within Teacher	L(K-1)	$S^2(poc/t) + J\ S^2(po/t) + IJ\ S^2(p/t)$
Error	Periods x Observer Within Teacher	L(K-1)(I-1)	$S^2(poc/t) + J\ S^2(po/t)$
True	Periods x Category Within Teacher	L(K-1)(J-1)	$S^2(poc/t) + I\ S^2(pc/t)$
Error	Periods x Observer x Category Within Teacher	L(I-1)(J-1)(K-1)	$S^2(poc/t)$

*Fixed variate

Weighted and Unweighted Scoring

The question of whether or not the verbal behavior categories from which the n Dom and n Intg scores were determined should be differentially weighted was raised in a previous chapter. It seemed reasonable to the investigators to assume that certain of the verbal behaviors were more potent contributors to the defined concepts of n Dom and n Intg than others. It was felt that category one, "accepts feeling," and category two, "praises or encourages," were stronger contributors to a climate of indirect influence than were categories three and four. It was also believed that category nine, "justifying authority," and eight, "criticizing," were more powerful in setting a tone of direct influence than categories six and seven. Because it might also be argued that these categories which were thought to be the most potent might be overpowered by the more frequent occurrence of the other categories, the investigators decided that the assumption should first be tested before using weighted rather than unweighted scores.

Scores for n Dom and n Intg were first obtained in the manner already described. For each visit made by an observer, the number of observations per hundred were determined for each category per five minute observation, and the values for each period were summed to give a total score for n Dom and n Intg for each observation. First and second quarter scores were the sum of the five observations made during a quarter; the total score for each subject was the total of both quarters. The quarter scores were used in testing the effects of the placement variables on these scores. The overall score was used to test the relationship with admissions data, the predictor variables in this study.

The technique of reciprocal averages was used to derive the weighted scores. The scaling method is described by Mosier (1942) and Mitzel & Hoyt (1954) as one which, beginning with a priori weights, assigns adjusted weights to each item response by an iterative process such that the contribution of each response to the total score is maximized. The procedure assumes that the variable considered can be represented by the sum of the item responses and that the majority of the items have been assigned a priori weights which were scaled in the appropriate direction.

Analyses made for this study were made by the ERA 1103 Univac Scientific Computer. The program available requires that the item input be in the form of single digit numbers having values from one to seven. To make use of this program it was necessary to translate the raw scores to a seven point scale. Since each verbal behavior category was treated as an item, each category score was coded one to seven. To simplify the coding system, the investigators decided that one coding system would be used for all of the categories. To estimate the distribution of raw scores of the eight categories for each of the ten observations for the thirty-six subjects, a sample of observations was drawn. The distribution was studied to suggest how the total range of score values could be divided into intervals which could then be coded using the seven point scale. In order to keep the data in their original form, weighted by the frequency of occurrence, the width of the intervals were made so that an equal percentage of the raw scores would fall within each interval. Because the cutting points for the intervals were made on a sample, the percentages were only approximate as seen in the tabulation of the percent of the total number of observations assigned each of the a priori scale values.

a priori scale	percent of total frequency
1	9
2	13
3	16
4	14
5	20
6	20
7	8
	100

Using the intervals established, the raw scores of each of the eight verbal behavior categories for each visit were coded using the seven point scale and punched into cards. Using the coded scale values as the a priori weights, adjusted weights and total scores for n Dom and n Intg were calculated for each of the ten observations separately. The weighted scores from each observation were then totaled to give n Dom and n Intg scores for the first and second quarter and for the two quarters combined.

Table 3.6 presents a comparison of the a priori and reciprocal average weights. From the distribution of responses for each verbal behavior category the mean weight was computed for each category using the a priori and reciprocal average weights. If the optimum weighting for each category was related to the frequency with which the behaviors were observed, then the two sets of weights should be equal or nearly so. The wide divergence of the weights, especially for categories one and six indicates that the assumption made by the investigators regarding the differential potency of verbal behaviors in the categories is supported.

Table 3.6

Mean A Priori and Reciprocal Average Weights for each Category of Verbal Behavior Comprising n Dom and n Intg Scores

n Dom			n Intg		
Category	Mean A Priori Weight	Mean RA Weight	Category	Mean A Priori Weight	Mean RA Weight
6	4.28	4.01	1	1.24	5.47
7	4.55	4.06	2	1.87	3.71
8	2.50	4.36	3	3.49	3.42
9	1.08	6.90	4	4.71	3.50

Table 3.7 presents a further comparison of the unweighted and RA scores. As would be expected in a technique which is designed to establish optimum weights, the standard deviation of the RA weights was considerably larger than that of the unweighted scores. It is also interesting to note that the correlation between the unweighted and RA scores is considerably higher for n Intg than for n Dom. Further detailed study of these data will be needed before the reason for the difference could be explained.

Table 3.7

Mean, Standard Deviation of Unweighted and Reciprocal Average Scores and Their Correlation for n Dom and n Intg.

	n Dom		**n Intg**	
	Unweighted	**RA**	**Unweighted**	**RA**
X	**5.8**	**8.7**	**3.4**	**9.2**
sd	**.91**	**1.40**	**.85**	**1.68**
r	**.32**		**.76**	

Chapter IV

The Relation of Observed Behaviors to Classroom Placement Variables

The previous chapter has presented the variables which were used in this study to describe the classroom behaviors of student teachers. For each of the variables descriptive statistical data were presented to characterize the central tendency of the observations, their variability and their interrelationships. In this chapter the statistical design which was used to identify and control classroom placement variables is presented, and each of the observational variables is analyzed according to the model.

The Cross-Over Design

Chapter 2 presented the sources of variability that were believed important in considering the approach to the three general questions with which the study was concerned. In addition the discussion identified those factors which could be varied or partitioned so that the extent of their influence could be extricated, and those whose influence could not be identified or estimated. These latter factors were controled through randomization procedures so that they would not bias the results.

The experimental factors were the following:

1. Preference for level of teaching: upper(4, 5, 6) grades or lower (1, 2, 3) grades
2. Experience in student teaching: First quarter or second quarter
3. Grade Level: upper (4, 5, 6) or lower (1, 2, 3)
4. Order of experience: upper followed by lower grade, or lower grade followed by upper

Figure 2.1 presented in Chapter 2 illustrates the arrangement of the four factors. Subjects eligible for placement in the experiment were first divided into two groups according to their stated grade level preference, those who preferred to teach upper grades (4, 5, 6) designated prefer U, and those who preferred to teach lower grades (1, 2, 3) designated prefer L. Each of these preference groups was randomly assigned to one of the two orders of experience. One-half of each preference group received its stated choice the first quarter. The other did not. The model specifying any individual observation can be stated thus:

> X(ijkl) is the observation score for the lth student teacher with the ith preference, teaching at the jth level, assigned the kth order so that

i = 1: prefer upper
i = 2: prefer lower
j = 1: upper grade
j = 2: lower grade
k = 1: upper-lower order
k = 2: lower-upper order
l = 1, 2, 3 . . . l identifies the student teacher

It is assumed that X(ijkl) are random variables, normally distributed with the following specifications.

$$
\begin{aligned}
&E[X(ijkl)] = W(ijk) \\
&\text{Variance } [X(ijkl)] = \sigma^2 \\
&\text{Covariance } [X(ijkl), X(i^1, j^1, k^1, l^1)] = \rho\sigma^2 = \\
&\text{for } i = i^1,\ j = j^1,\ k \neq k^1, \text{ otherwise zero.}
\end{aligned}
$$

W(ijkl) is further defined as a linear function of fixed parameters which are representative of the various factors affecting a score. Thus

$$
\begin{aligned}
W(ijk) = {} & M + P(i) + L(j) + O(k) + PL(ij) + \\
& LO(jk) + PLO(ijk)
\end{aligned}
$$

The analysis of variance for this cross-over design is presented in Table 4.1 and the calculation formulas for computing the sums of squares is given in Table 4.1a. This analysis follows the pattern given by Cochran & Cox (1950) and Stanley (1955).

Table 4.1

Analysis of Variance Model for Counterbalanced Observations of Student Teachers at Both Upper and Lower Elementary Grades over a Two Quarter Period Stratified According to Grade Level Preference

SV	df	SS
Set 1		
L (j) (level)	L - 1	C - Z
LO (jk) (Quarter)	(L - 1) (O - 1)	E - B - C - Z
PL (ij)	(P - 1) (L - 1)	F - C - A + Z
PLO (ijk)	(P - 1) (L - 1) (O - 1)	G - E - D - F + A + B + C - Z
Error (Set 1)	PO (L - 1) (K - 1)	T - H - G + D
Total (Set 1)	POK (L - 1)	T - H
Set 2		
M	1	Z
P (i) (preference)	P - 1	A - Z
O (k) (order)	O - 1	B - Z
PO (ik)	(P - 1) (O - 1)	D - B - A + Z
Error (Set 2)	PO (K - 1)	H - D
Total (Set 2)	POK - 1	H

Table 4.1a

Computation Formulas for Sums of Squares for Analysis of Variance Model

$X^2 (\ldots.) \;/\; PLOK = Z$

$E_{j} X^2 (p\ldots) \;/\; LOK = A$

$E_{o} X^2 (..o.) \;/\; PLK = B$

$E_{l} X^2 (.l..) \;/\; POK = C$

$E_{po} X^2 (p.o.) \;/\; LK = D$

$E_{lo} X^2 (.lo.) \;/\; PK = E$

$E_{pl} X^2 (pl..) \;/\; OK = F$

$E_{plo} X^2 (plo.) \;/\; K = G$

$E_{pok} X^2 (p.ok) \;/\; L = H$

$E_{plok} X^2 (plok) = T$

The following two sections present the analysis of data from the OScAR III and from the analysis of verbal behaviors. Each of the tables gives the results of the tests of seven hypotheses, three main effects, which are, preference, order, and level, and the four designated interactions. In the discussion of each of the tables the results are announced only for those tests where significant differences were found.

Analysis of OScAR III Scales

Table 4.2 presents the analysis of the Teacher-Individual Activities (TIA) scale. One significant F value is noted indicating that TIA scores were different at the two different grade levels. When student teachers were in the upper grade level the mean TIA score was 273.1 while the mean in the lower grade level was 258.6. This result indicates that when in the upper grades student teachers were more often observed giving positive attention and encouragement to individual pupils than when they were in the lower grades. In using the scale items to interpret this result it must be kept in mind that some items of the scale were scored negatively.

Table 4.2

Analysis of Variance of Teacher-Individual Activities for Counterbalanced Observations of Student Teachers at Both Upper and Lower Elementary Grades over a Two Quarter Period Stratified According to Grade Level Preference.

SV	df	SS	MS	F
Set 1				
L (j) (levels)	1	3574.250	3574.250	4.24*
LO (jk)	1	642.368	642.368	>1
PL (ij)	1	148.364	148.364	>1
PLO (ijk)	1	8.911	8.911	>1
Error (Set 1)	30	25286.607	842.8869	
Total (Set 1)	34	29660.500		
Set 2				
P (i) (preferences)	1	2.795	2.795	>1
O (k) (order)	1	420.015	420.015	>1
PO (ik)	1	525.690	525.690	>1
Error (Set 2)	30	28653.515	955.1171	
Total (Set 2)	33	29602.015		

*significant at .05 level
**significant at .01 level

Table 4.3 gives the results of the tests of hypotheses regarding the Teacher Class Activities (TCA) scale. The only significant difference found involved the level by order, B(lo), interaction. As indicated in discussions of the cross-over model, the level by order interaction is confounded with the quarter effect and is most directly interpreted as an effect due to the quarter of experience. The four mean scores which the test indicates are different are the following:

Order	Level	
	upper	lower
Upper-lower	243	264
Lower-upper	265	254

It can be seen from these means that the two groups having a different order of experience in the two grade levels had significantly higher mean scores in their second quarter, 265 and 264, than in the first. These results indicate that during their second quarter student teachers provided a wider variety of teacher-class activities than they did their first quarter of student teaching. It should also be noted that the increase from first to second quarter was greatest for the upper-lower group.

Table 4.3

Analysis of Variance of Teacher Class Activities for Counterbalanced Observations of Student Teachers at Both Upper and Lower Elementary Grades over a Two Quarter Period Stratified According to Grade Level Preference.

SV	df	SS	MS	F
Set 1				
L (j) (levels)	1	395.529	395.529	>1
LO (jk)	1	4130.883	4130.883	6.069*
PL (ij)	1	18.354	18.354	>1
PLO (ijk)	1	806.908	806.908	1.185
Error (Set 1)	30	20420.326	680.6775	
Total (Set 1)	34	25772.000		
Set 2				
P (i) (preferences)	1	1674.037	1674.037	2.132
O (k) (order)	1	636.235	636.235	>1
PO (ik)	1	260.375	260.375	>1
Error (Set 2)	30	23560.235	785.3411	
Total (Set 2)	33	26130.882		

*significant at .05 level
**significant at .01 level

The tests of hypotheses concerning Teacher Group Activities (TGA) are given in Table 4.4. None of the experimental factors was found to have a significant effect on this observation variable.

One significant F value is reported in Table 4.5. It indicates that there were differences in Pupil Class Activities (PCA) when student teachers were teaching at different grade levels. The mean score value for (PCA) at the upper grade level was 244 and 260 at the lower grade level. The variety of activities in which pupils led the class were larger when student teachers were in the lower grades than when they were in the upper grades.

Table 4.6 presents the cross-over analysis for Pupil Disruptive Behaviors (PDB). Tests of hypotheses found no significant F values indicating that none of the experimental factors affected the observed pupil disruptive behaviors beyond chance expectations.

In the analysis of Classroom Variety of Materials (CVM), Table 4.7, one factor was found to affect the scale scores significantly. This factor occurred "within individuals" as was true for the other results thus far announced. The mean CVM score for student teachers when in the upper grades was 264 while the mean when they were in the lower grades was 247. Such a difference indicates that student teachers used a wider variety of teacher and pupil materials when in the upper grades.

Table 4.4

Analysis of Variance of Teacher Group Activities for Counterbalanced Observations of Student Teachers at Both Upper and Lower Elementary Grades over a Two Quarter Period Stratified According to Grade Level Preference.

sv	df	SS	MS	F
Set 1				
L (i) (levels)	1	1948.470	1948.470	4.073
LO (jk)	1	147.060	147.060	>1
PL (ij)	1	95.302	95.302	>1
PLO (ijk)	1	398.896	398.896	>1
Error (Set 1)	30	14350.272	478.342	
Total (Set 1)	34	16940.000		
Set 2				
P (i) (preferences)	1	137.655	137.655	>1
O (k) (order)	1	226.117	226.117	>1
PO (ik)	1	245.655	245.655	>1
Error (Set 2)	30	14064.455	468.815	
Total (Set 2)	33	14673.882		

*significant at .05 level
**significant at .01 level

Table 4.5

Analysis of Variance of Pupil Class Activities for Counterbalanced Observations of Student Teachers at Both Upper and Lower Elementary Grades over a Two Quarter Period Stratified According to Grade Level Preference.

sv	df	SS	MS	F
Set 1				
L (j) (levels	1	4368.015	4368.015	6.595*
LO (jk)	1	1832.485	1832.485	2.767
PL (ij)	1	86.482	86.482	>1
PLO (ijk)	1	8.646	8.646	>1
Error (Set 1)	30	19870.872	662.3624	
Total (Set 1)	34	26166.500		
Set 2				
P (i) (preferences)	1	92.123	92.123	>1
O (k) (order)	1	2886.015	2886.015	2.980
PO (ik)	1	1594.391	1594.391	1.646
Error (Set 2)	30	29052.780	968.426	
Total (Set 2)	33	33625.309		

*significant at .05 level
**significant at .01 level

Table 4.6

Analysis of Variance of Pupil Disruptive Behavior for Counterbalanced Observations of Student Teachers at Both Upper and Lower Elementary Grades over a Two Quarter Period Stratified According to Grade Level Preference.

SV	df	SS	MS	F
Set 1				
L (j) (levels)	1	3234.720	3234.720	3.938
LO (jk)	1	22.368	22.368	>1
PL (ij)	1	2097.288	2097.288	2.553
PLO (ijk)	1	0.912	0.912	>1
Error (Set 1)	30	24641.212	821.3737	
Total (Set 1)	34	29996.500		
Set 2				
P (i) (preferences)	1	333.818	333.818	>1
O (k) (order)	1	.368	.368	>1
PO (ik)	1	.003	.003	>1
Error (Set 2)	30	37586.061	1252.8687	
Total (Set 2)	33	37920.250		

*significant at .05 level
**significant at .01 level

Table 4.7

Analysis of Variance of Classroom, Variety of Materials for Counterbalanced Observations of Student Teachers at Both Upper and Lower Elementary Grades over a Two Quarter Period Stratified According to Grade Level Preference.

SV	df	SS	MS	F
Set 1				
L (j) (levels)	1	5084.471	5084.471	5.127*
LO (jk)	1	1927.117	1927.117	>1
PL (ij)	1	105.775	105.775	>1
PLO (ijk)	1	7.464	7.464	>1
Error (Set 1)	30	29749.173	991.6391	
Total (Set 1)	34	36874.000		
Set 2				
P (i) (preferences)	1	9.364	9.364	>1
O (k) (order)	1	933.883	933.883	1.328
PO (ik)	1	1528.332	1528.332	2.173
Error (Set 2)	30	21097.539	703.2513	
Total (Set 2)	33	23569.118		

*significant at .05 level
**significant at .01 level

Table 4.8 shows several significant results when the experimental effects on Classroom Emotional Climate (CEC) were analyzed. One main effect was significant. Mean CEC scores were significantly higher when student teachers were in upper grades. The preference by level interaction, B (pl), was also significant. The four means involved are the following:

Preference	Level	
	upper	lower
Prefer U	261	233
Prefer L	250	251

The significance of this interaction, as well as the difference between the levels, can be attributed to the prefer upper group, for the prefer lower group showed no change in CEC scores at the two different levels. The prefer upper group had significantly better CEC scores when they were in the upper grades than when they were in the lower. The G (plo) interaction was also significant. The eight means are given below:

Preference	Order	Level	
		upper	lower
upper	upper-lower	261	231
	lower-upper	261	235
lower	upper-lower	252	251
	lower-upper	248	250

The explanation for this interaction is found principally in the prefer upper group. In this group the CEC scores were equal in the upper grades regardless of the order of experience, but CEC scores were highest in the lower grades for those who taught first in the lower grades. An alternative explanation which is possible would consider G (plo) as a preference by quarter interaction (lo = quarter), but an examination of the means does not indicate that such an explanation is plausible, especially in veiw of the lack of "level by order" (quarter) effect. It is more reasonable to conclude that G (plo) interaction was an extension of the B (pl) interaction to the order dimension.

The analysis of the final OScAR III scale Classroom Verbal Emphasis (CVE), is presented in Table 4.9. Two significant main effects are indicated, differences between levels and between preferences. The significant difference between preference groups indicates a between groups difference. Those student teachers who preferred the lower grades had a higher mean CVE score ($\overline{X} = 256$) than those who preferred the upper grades ($\overline{X} = 240$). The level effect represents a "within individuals" difference. Student teachers regardless of preference had higher CVE scores when they were in the upper grades ($\overline{X} = 256$) than when they were in the lower grades ($\overline{X} = 244$).

Table 4.8

Analysis of Variance of Classroom Emotional Climate for Counterbalanced Observations of Student Teachers at Both Upper and Lower Elementary Grades over a Two Quarter Period Stratified According to Grade Level Preference.

sv	df	SS	MS	F
Set 1				
L (j) (levels)	1	1487.117	1487.117	4.946*
LO (jk)	1	46.118	46.118	>1
PL (ij)	1	3081.766	3081.766	10.250**
PLO (ijk)	1	14015.491	14015.491	46.617**
Error (Set 1)	30	9019.508	300.6502	
Total (Set 1)	34	27650.000		
Set 2				
P (i) (preferences)	1	192.299	192.299	>1
O (k) (order)	1	9.941	9.941	>1
PO (ik)	1	66.669	66.669	>1
Error (Set 2)	30	41777.326	1392.5775	
Total (Set 2)	33	42046.235		

*significant at .05 level
**significant at .01 level

Table 4.9

Analysis of Variance of Classroom Verbal Emphasis for Counterbalanced Observations of Student Teachers at Both Upper and Lower Elementary Grades over a Two Quarter Period Stratified According to Grade Level Preference.

sv	df	SS	MS	F
Set 1				
L (j) (levels)	1	2508.367	2508.267	4.635*
LO (jk)	1	492.486	492.486	>1
PL (ij)	1	36.368	36.368	>1
PLO (ijk)	1	7.219	7.219	>1
Error (Set 1)	30	16236.060	541.2020	
Total (Set 1)	34	19280.500		
Set 2				
P (i) (preferences)	1	3785.847	3875.847	4.3012*
O (k) (order)	1	14.132	14.132	>1
PO (ik)	1	1430.739	1430.739	>1
Error (Set 2)	30	26405.061	880.1687	
Total (Set 2)	33	31635.779		

*significant at .05 level
**significant at .01 level

Analysis of Verbal Behaviors

The influence of the factors considered in the cross-over design upon Teacher Verbal Output (TVO) is shown by the data in Table 4.10. The one significant F value indicates that the mean score representing the amount of teacher talk was higher when student teachers were in the upper grades than when they were in the lower grades. The mean for upper grades was 275 and the mean for lower grades was 255.

Table 4.10

Analysis of Variance of Teacher Verbal Output for Counterbalanced Observations of Student Teachers at Both Upper and Lower Elementary Grades over a Two Quarter Period Stratified According to Grade Level Preference.

sv	df	SS	MS	F
Set 1				
L (j) (levels)	1	6800.000	6800.000	10.228**
LO (jk)	1	1487.148	1487.148	2.237
PL (ij)	1	545.030	545.030	>1
PLO (ijk)	1	640.882	640.882	>1
Error (Set 1)	30	19445.940	664.8646	
Total (Set 1)	34	28919.000		
Set 2				
P (i) (preferences)	1	3.426	3.426	>1
O (k) (order)	1	1640.500	1640.500	1.529
PO (ik)	1	281.530	281.530	>1
Error (Set 2)	30	32178.485	1072.616	
Total (Set 2)	33	34103.941		

*significant at .05 level
**significant at .01 level

Tables 4.11 and 4.12 present the analysis of Need Integration (n Intg) and Need Dominance (n Dom) scores computed from the sum of the unweighted item frequencies. Table 4.11 indicates that the placement factors had no influence on the n Intg scores thus derived. Table 4.12 indicates that two interactions are significant when n Dom scores were considered. The four means are the following:

order	quarter	
	1st	2nd
upper-lower	6.73	4.80
lower-upper	6.88	4.72

Table 4.11

Analysis of Variance of Need Integration Scores for Counterbalanced Observations of Student Teachers at Both Upper and Lower Elementary Grades over a Two Quarter Period Stratified According to Grade Level Preference

sv	df	SS	MS	F
Set 1				
L (j) (levels)	1	.0777	.0777	>1
LO (jk)	1	24.3602	24.3602	6.577
PL (ij)	1	.0023	.0023	>1
POL (ijk)	1	.1329	.1329	>1
Error (Set 1)	30	11.1119	3.7039	
Total (Set 1)	34	35.6850		
Set 2				
P (i) (preferences)	1	.1415	.1415	>1
O (k) (order)	1	.1236	.1236	>1
PO (ik)	1	.1599	.1599	>1
Error (Set 2)	30	51.3881	1.7129	
Total (Set 2)	33	51.8131		

*significant at .05 level
**significant at .01 level

Table 4.12

Analysis of Variance of Need Dominance Scores for Counterbalanced Observations of Student Teachers at Both Upper and Lower Elementary Grades over a Two Quarter Period Stratified According to Grade Level Preference.

sv	df	SS	MS	F
Set 1				
L (j) (levels)	1	1.5001	1.5001	>1
LO (jk)	1	58.9248	58.9248	49.4584**
PL (ij)	1	3.1852	3.1852	2.6734
PLO (ijk)	1	5.6207	5.6207	4.7177*
Error (Set 1)	30	35.7442	1.1914	
Total (Set 1)	34	104.9750		
Set 2				
P (i) (preferences)	1	1.1501	1.1501	>1
O (k) (order)	1	.8272	.8272	>1
PO (ik)	1	2.3399	2.3399	>1
Error (Set 2)	30	57.7777	1.9259	
Total (Set 2)	33	62.0949		

*significant at .05 level
**significant at .01 level

These means indicate that n Dom scores were significantly lower for all student teachers during the second quarter of student teaching as compared with the first quarter. The means for the preference by level by order (preference by quarter) interaction are the following:

Preference	Order	Quarter	
		1st	2nd
upper	upper-lower	5.36	5.16
	lower-upper	7.08	5.11
lower	upper-lower	6.91	4.51
	lower-upper	6.78	4.60

The above n Dom means shows one deviation from the general quarter effect shown in the previous set of means thus accounting for the interaction. Those women who preferred the upper grades and who taught first at the preferred level had n Dom scores that differed little from the first to the second quarter. Their first quarter scores more nearly resembled the second quarter scores of the other group than the first quarter scores of that group.

Tables 4.13 and 4.14 present the analysis of n Intg (RA) and n Dom (RA) scores. These data represent the same initial observations as presented in Tables 4.11 and 4.12. The RA designation indicates that the reciprocal averages technique was employed to weight differentially the various item responses according to their contribution to the total score. The effect of using this scaling technique on these data has been described in the previous chapter.

Table 4.13 indicates that for n Intg (RA) the interaction of level by order (Quarter) was significant. The mean n Intg (RA) scores are the following:

order	level	
	upper	lower
upper-lower	33	58
lower-upper	56	38

The means indicate that the scores for n Intg (RA) were significantly higher the second quarter of student teaching than the first. Table 4.14 presents the tests of significance for n Dom (RA) scores. These tests indicate that three F values were significant. All three refer to main effects if level by order is considered a quarter difference. Level and level by order are within individual differences, but order differences are between individual differences. The means for these three tests follow .

Table 4.13

Analysis of Variance of Need Integration (RA) Scores for Counterbalanced Observations of Student Teachers at Both Upper and Lower Elementary Grades over a Two Quarter Period Stratified According to Grade Level Preference.

sv	df	SS	MS	F
Set 1				
L (j) (levels)	1	194.4853	194.4853	2.087
LO (jk)	1	7687.1911	7687.1911	82.4864**
PL (ij)	1	13.7039	13.7039	>1
PLO (ijk)	1	143.3166	143.3166	>1
Error (Set 1)	30	2795.8031	93.1934	
Total (Set 1)	34	10834.5000		
Set 2				
P (i) (preferences)	1	27.9225	27.9225	>1
O (k) (order)	1	51.1912	51.1912	>1
PO (ik)	1	240.9980	240.9980	1.712
Error (Set 2)	30	4223.1972	140.7732	
Total (Set 2)	33	4543.3089		

*significant at .05 level
**significant at .01 level

Table 4.14

Analysis of Variance of Need Dominance (RA) Scores for Counterbalanced Observations of Student Teachers at Both Upper and Lower Elementary Grades over a Two Quarter Period Stratified According to Grade Level Preference.

sv	df	SS	MS	F
Set 1				
L (j) (levels)	1	986.4853	986.4853	11.2750**
LO (jk)	1	718.2498	718.2498	8.2092**
PL (ij)	1	19.5223	19.5223	>1
PLO (ijk)	1	83.4544	83.4544	>1
Error (Set 1)	30	2624.7882	87.4929	
Total (Set 1)	34	4432.5000		
Set 2				
P (i) (preferences)	1	7.5436	7.5436	>1
O (k) (order)	1	440.1324	440.1324	4.5609*
PO (ik)	1	.6025	.6025	>1
Error (Set 2)	30	2895.0304	96.5010	
Total (Set 2)	33	3343.3089		

*significant at .05 level
**significant at .01 level

order	level		
	upper	lower	$\bar{X}$
upper-lower	40.5	41.6	41.1
lower-upper	39.1	53.2	46.2
$\bar{X}$	39.8	47.4	

The table of means indicates that n Dom (RA) scores were significantly higher for the group of students who first taught in the lower grades then moved to the upper grades for the second quarter of experience. It also indicates that for all students the n Dom (RA) scores were higher in the lower grades than in the upper grades. An inspection of the four level by order means suggests that the difference cannot be interpreted as a between quarters difference. The mean score (53.2) of the group who had the lower-upper order when they were at the lower level (first quarter) is responsible for both significant main effects and the interaction.

A comparison of the results of the tests of significance using the unweighted and weighted (RA) scores indicates several differences which can be attributed to the difference in scores. Only one of the differences was noted with n Intg scores. The level by order interaction (quarter effect) was found to be significant when the weighted scores (RA) were used. For need Dominance two additional significant effects were noted when the weighted scores were used, and one significant interaction dropped out.

From the inception of the study, one of the purposes for establishing the experimental design was so that the effects of the placement variables on the observed behaviors could be identified. If it were necessary to compensate for placement effects, adjustments could be made in the observed variables before they were used as criterion variables in subsequent analyses. The scores used as criterion variables and thus of primary concern were n Intg (RA) and n Dom (RA). The analyses in this chapter found that n Intg (RA) scores were in general higher the second quarter than the first. Since these differences were "within individuals," they do not affect "between individuals" comparisons. The n Dom (RA) scores, however, did show one "between individuals" difference. Those student teachers who taught in the lower then in the upper grades had higher n Dom (RA) scores than those who had the reverse order assignment. Any interpretation of this order effect is complicated by the level by order interaction which was also found to be significant. Because the effects of the placement variables on n Dom (RA) scores could not be interpreted as a straight forward "between individuals" difference, the investigators decided that no adjustments would be made in n Dom (RA) scores before they were used as criterion scores in the regression analysis.

Chapter V

The Relation of Predictor Variables to Observed Behaviors

One of the purposes of this study was to determine the relationship between data collected at admission and the observed classroom behaviors of student teachers. There are a number of methods of examining the relationship among the variables singly or of considering them collectively. Since the most probable application of the results of these analyses would be to predict teacher behavior from admissions data, it was decided that multiple regression procedures would be appropriate. Such analyses test whether the variables are predictive when used in a linear combination. Because the relations among the variables are useful descriptive data, this chapter first presents the intercorrelation analysis. These data examine the extent of the relationship among the predictor variates as well as their relationship to the criterion. Following the descriptions of the bivariate relationships, the multivariate analyses are presented.

Relationships Among the Variables

In Table 5.1 the correlations among the predictor variables is given as is the correlation of each predictor with the appropriate criterion variables. Only five of the thirty-two correlations were larger than expected by chance. Three of the five were among predictor variables, and these r's are of the order that would be expected on the basis of previous studies. High school rank (HSR) and the Miller Analogies Test (MAT) correlate significantly with sophomore honor point ratio (HPR), and the MAT correlates with the MTAI beyond chance expectations. None of the other correlations among the predictors were significant.

Table 5.1

Intercorrelations Among Predictor Variables and Among Predictor and Criterion Variables

	Predictors						Criterion	
	HPR	MAT	MTAI	Bowers	n Dom(c)	n Intg(c)	n Dom (RA)	n Intg (RA)
HSR	.50**	.27	-.03	-.06	.10	-.02	-.09	.11
HPR		.39*	-.01	-.04	.02	.19	-.19	.38*
MAT			.35*	.04	.05	-.16	-.04	.31
MTAI				-.13	.04	.15	-.18	.44*
Bowers					-.06	-.10	-.02	.06
n Dom (c)							.01	
n Intg (c)								.18

(c) = counselors' rating

When the relationship among the predictors and the criterion variable for need Dominance, n Dom (RA), were examined no significant correlations were found. Noteworthy, however, is the fact that all of the predictors correlate negatively with n Dom (RA) except counselor's predictions of n Dom which were positive though near zero.

In contrast when n Intg (RA) was the criterion variable all of the correlations were positive, two reaching significance. Sophomore HPR and MTAI scores were related to classroom observations of n Intg as indicated by correlations of .38 and .44.

The intercorrelations suggested that a linear combination of predictors might be successful in predicting the criterion variables. The chain-like relation between HSR and HPR, HPR and MAT, MAT and MTAI with low correlations among other pairs increased the possibility of a significant linear combination. The following section presents the statistical model and the results of the analyses.

Multiple Regression Analysis

The Statistical Model. The aim of this analysis is to predict certain aspects of classroom performance, Y, of elementary student teachers from a knowledge of admissions data, the six predictor variables. The sample of student teachers in the study is considered to be a random sample from a hypothetical population of elementary student teachers. For this population the multiple regression model, after the manner of Walker & Lev (1953), is

$$E(Y) = \alpha + \beta_1 X_1 + \beta_2 X_2, \ldots, + \beta_6 X_6$$

where E(Y) is the average score for all individuals in the population with a particular set of scores on the predictor variables. The population regression coefficients, the α and the β's, are unknown weighting constants appropriate for this population. The data from the study provide sample estimates, A, b, b_2, . . ., b_6, of the population regression coefficients. These sample estimates are used to obtain an equation for predicting Y scores. The equation for obtaining these predicted Y scores is

$$Y_i = A + b_1 X_{1i} + b_2 X_{2i} \ldots, + b_6 X_{6i}$$

where

Y_i denotes the predicted n Dom (RA) or n Intg (RA) score for the ith student

X_{1i} denotes the HSR for the ith student

X_{2i} denotes the sophomore HPR for the ith student

X_{3i} denotes the MAT score for the ith student

X_{4i} denotes the MTAI score for the ith student

X_{5i} denotes the Bowers Inventory score for the ith student

X_{6i} denotes the counselor's predictors of n Dom or n Intg for the ith student

Several tests of significance about the population model and the population regression coefficients can be performed to indicate the usefulness of this linear combination of variables for predicting classroom performance. In order to make these tests of significance several assumptions are required. In the population the Y's for each set of X's are assumed to be normally distributed around their expected value, and all such Y distributions are assumed to have the same variance. It is also assumed that $X_1, X_2, \ldots, X_6$ are fixed variables and that predictions can be made only for those sets of X's which actually appear in the sample.

To find the b's, the weighting for each predictor variable, the set of equations may be written in matrix form, [A][B] = [G] where [A] is the matrix of sums and cross products among the predictor variables expressed in deviation form and where [G] is the matrix of the sums of cross products between the predicted variable, Y, and the predictor variables, $X_1, X_2, \ldots, X_6$ also in deviation form. [B] is the matrix of the b's, the partial regression coefficients, the unknowns in this set of equations. The solution of this set of equations is arrived at by finding the inverse, $[A]^{-1}$, by pivotal condensation and finding the b's from $[B] = [A]^{-1}[G]$.

Once the partial regression coefficients are obtained the multiple correlation, R, may be found from the formula

$$R = \sqrt{\frac{B'G}{\Sigma y^2}}$$

The next step is to test the significance of the multiple correlation coefficient by means of an F test, the test statistic being

$$F = \frac{R^2}{1 - R^2} \times \frac{N - M - 1}{M}$$

where M is the number of predictor variables. The critical values are obtained from the F table with M and N - M - 1 degrees of freedom.

When the multiple correlation is found to be significantly different from zero, the contribution of the individual partial regression coefficients is considered. The significance of each of the b's is determined by the following F test.

$$F = \frac{b_j^2}{\frac{1 - R^2}{N - M - 1} \times a^{jj}}$$

with 1 and N - M - 1 degrees of freedom where a^{jj} is the appropriate term in the diagonal of the inverse matrix. Of the predictor variables for which the b's are judged to be non-significant, one may be removed from the equation, the new partial regression coefficients being computed from the formula

$$b_j^* = b_j - \frac{b_d a^{dj}}{a^{dd}}$$

where d is the subscript of the variable eliminated at each stage.

The new inverse is then computed using the formula

$$a^{*ij} = a^{ij} - \frac{a^{id} a^{dj}}{a^{dd}}$$

The significance of the new regression weights can be tested by

$$F = \frac{[b_j^*]^2}{\frac{1 - R^{*2}}{N - M - 1} \times a^{*jj}}$$

with 1 and N - M* - 1 degrees of freedom where M* = M - 1.

This procedure is repeated as long as there is no significant reduction in R. The significance of the difference between R^2 and R^{*2} can be tested by the F test,

$$F = \frac{[R^2 - R^{*2}]\,[N - M - 1]}{[1 - R^2]\,[M - M^*]}$$

with M - M* and N - M - 1 degrees of freedom.

Finally the multiple regression equation is written using only the variables retained after the final test of the partial regression coefficients.

Two multiple regression equations were computed, one for each of the two measures of observed teacher behaviors, n Dom (RA) and n Intg (RA).

Predicting n Dom

The first regression equation was computed using n Dom (RA) as the predicted, or dependent, variable. The obtained value for R was .3077. (The actual computations are contained in Appendix C.) An F test was computed to test the hypothesis that the relationship of the combined predictor variables to the criterion variable could be accounted for by chance, and this hypothesis was accepted at the .05 level of significance. The admissions data were not significantly related to classroom observations of n Dom.

Predicting n Intg

The second regression equation contained n Intg (RA) as the dependent variable. The obtained value of R was .5987. (The actual computations are contained in Appendix C.) The null hypothesis that the relationship between the independent and dependent variables could be accounted for by chance was rejected at the .05 level of significance, and it was concluded that a significant relationship between the linear combination of independent variables and the dependent variable, (n Intg (RA)), did in fact exist.

However, tests of the individual contributions of each of the predictor variables disclosed that only two of the variables, sophomore HPR and MTAI score were contributing significantly to the regression equation. The four nonsignificant variables were deleted and a new regression equation was computed. The adjusted value of R, using only the two variables, was .5775, a value significantly larger than chance. An individual test of the two variables showed that each made a significant contribution to the new regression equation.

The possibility that the four deleted variables had, in combination with the remaining two, made a significant contribution to the regression equation was tested by comparing the first multiple correlation, (R^2), with the second, (R^{*2}). The F test revealed no significant difference between the two equations. It was concluded that the four deleted variables did not contribute significantly to the multiple regression equation; thus the two variables, HPR and MTAI produced as high a relationship with the criteria as did all six predictors.

Chapter VI

The Comparisons of Ratings by Observers and by Supervisors

The third question which Chapter One posed for consideration concerned a comparison of the ratings made by research observers and those made by student teaching supervisors. The phrasing of the question took two forms.

1. Using the personality variables of concern in this study as the basis for comparison, what is the relationship of ratings of student teachers made by experienced student teaching supervisors using their established observational methods and the ratings derived from the observational techniques designed for this study and employed by the research observers?

2. Using certain teaching behaviors as the basis for comparison, what is the relationship of ratings of student teachers by student teaching supervisors using their established observational methods and the ratings made by research observers?

A comparison of these two questions suggests that the primary difference between them involves the basis for comparing the results of observations. In both cases the observational methods remained the same; research observers used techniques prescribed for the use of the OScAR III and the Minnesota System of Interaction Analysis, and the University supervisors used the observational methods that had been developed for the supervision of the elementary student teaching program. In the first question the bases for comparison were the personality variables, n Dom and n Intg, which were basic to other aspects of this study. In the second question certain classroom behaviors not thus far of concern in this study were the basis for comparison. In this chapter each of the questions under consideration is presented in a separate section.

Ratings of Personality Variables

The purpose of the first question posed in this chapter was to find, albeit in a gross sort of way, whether the variables, n Dom and n Intg, were sufficiently well defined that they were observable by observers and supervisors who had different backgrounds and who used different observational methods. If significant correlations were found between the judgments made by supervisors and the scores derived from observers application of the OScAR III and the Minnesota System of Interaction Analysis, then the relationship would be an indication, at least, that the same variables were being measured. In this study there was no way to find to what extent either the method of observation used by the supervisors and research observers or the background of experience of the two groups contributed to making the ratings similar or dissimilar. If no relationships were found it might be caused by lack of definition of the variables, the methods of observation, or the methods of measurement used. If significant relationships were found, it might be assumed that the variables had sufficient behavioral definition that they could be detected by the two different groups using different observational procedures and different scales of measurement.

Both the University supervisors and the research observers were given the definitional statements of n Dom and n Intg (see Chapter I) at the beginning of the study. Each group had an opportunity to learn the concepts as they were applied in this study, the observers through their training sessions, and the supervisors through the orientation sessions which were held with them in preparation for this study. The supervisors judgments were quantified by use of the rating scale included in Appendix A. At the close of the first and again at the end of the second quarter, supervisors were asked to summarize their judgments of the extent to which n Dom and n Intg were present in each subject by use of the rating scale. The n Dom and n Intg values representing the observers and their method of observation were those derived from the observational procedures described in earlier chapters.

Table 6.1 presents the correlations between the scores obtained from supervisors and observers for each quarter and for the total of the two quarters.

Table 6.1

The Correlation Between n Dom and n Intg Scores Obtained from Judgments of University Supervisors and Those Derived from Procedures Used by Research Observers

	n Dom		n Intg	
	Unweighted	RA	Unweighted	RA
Fall	.09	-.05	.42**	.32
Winter	.19	.38*	.32	.29
Total	.13	.11	.41**	.43**

*significant at .05 level
**significant at .01 level

In general there seemed to be little relationship between the n Dom scores which represented the observations of the two groups. During the winter quarter the correlation was significant when the scores for the research observers were weighted by the technique of reciprocal averages. Although the relationship between the unweighted scores also increased numerically neither of the coefficients was larger than could be expected by chance. When both quarters were considered together the relationship was not a significant one. When n Intg was considered, however, the picture was somewhat different. In general it can be said that there was an agreement beyond chance expectations in the order in which the subjects were ranked on n Intg by these two different procedures. Those coefficients which were not singificant were quite close to the level required for significance. It seemed that the RA weighting of research observers scores had little effect on the correlations of either variable.

It may be concluded that there was noticeable agreement between the University supervisors and the research observers on the degree to which n Intg was observed in student teachers. When n Dom was the variable under consideration no such agreement was found.

Ratings of Teaching Behaviors

The measuring scale used for answering the second question asked at the outset of this chapter was constructed especially for this study. The purpose of the instrument was to assess the extent to which two objectives of student teaching were observed: (1) the development of creativity, and (2) the achievement of harmonious interpersonal relations. In addition scales were used for rating the overall effectiveness in student teaching and the predicted success in teaching. Items were constructed to reflect the construct of creativity as outlined by Lowenfeld (1958) and the interpersonal relations objectives of the student teaching program at the University of Minnesota. Copies of the scales are included in Appendix A. Both supervisors and research observers used the same scales for quantifying their observations, and they completed their ratings after concluding the observations typically made: (1) the supervisor observing for purposes of evaluating the student teacher's progress and (2) the research observer for completing the OScAR III and the Minnesota System of Interaction Analysis. Because the same instrument was used by both supervisors and research observers, a more detailed comparison of their ratings is warranted. It is still true, however, that any differences noted must be attributed to a combination of the factors of background and experience and method of observation. In this study there was no way to extricate these two factors.

The comparison between supervisors and research observers was made in several ways. Statistics describing the distributions of ratings for the two groups are presented, as are the reliabilities of each set of ratings. The relationship between the two ratings is explored in two ways. First the correlations are given. Secondly, because each research observer saw each subject each quarter and the supervisors saw only those who were assigned to them at the beginning of the study, the "balanced" observations of the research observers were compared to the "unbalanced" observations of the supervisors. The analysis was based on the discrepancy of score differences between the composite of supervisors' ratings of each of the subjects and the mean score computed from the observations of all five of the research observers.

Table 6.2 presents the descriptive statistics, reliabilities, and the intercorrelations between the scores from supervisors and those from research observers. In the case of all three of the variables the mean scores for supervisors were higher and the standard deviations were larger. These results indicate that the observers tended to be more conservative in their use of the rating scale than did the supervisors, and the observers were, as a group, more homogeneous in their ratings.

The estimates of the reliabilities show some differences between supervisors and observers. These differences are largely accounted for by the methods used to estimate the reliability. Since the supervisors rated only those students assigned to them, the reliability coefficient is estimated by determining consistency with which each supervisor rated his student teachers and pooling the consistency of all supervisors.

Table 6.2

Data on Supervisors' and Observers' Ratings of Student Teachers on Four Variables Giving Mean, Standard Deviation, Reliabilities, the Correlation Between Ratings, and the Correlation Corrected for the Unreliability of Supervisors' and Observers' Ratings.

	Supervisors' Ratings			Observers' Ratings			Observer-Supervisor	
	$\bar{X}$	SD	REL	$\bar{X}$	SD	REL	r	Corrected r
Creativity & Resourcefulness	47.2	8.4	.80	42.4	5.4	.72	.62**	.82
Interpersonal Relations	28.5	5.5	.78	25.4	4.0	.73	.66**	.87
Success as a Student Teacher	3.6	.8	.77	3.3	.7	.71	.70**	.95

Cronbach has referred to this type of data as having "unmatched" observations, and in this case all variance among the judgments by any one supervisor is treated as error variance. The reliabilities which were computed for the observers were of the internal consistency type which assumed that observations were "matched," i.e., that all observers rated each student teacher. The reliability estimates for supervisors can be interpreted as the consistency, on the average, with which the supervisors rated the student teachers. The reliability coefficients for the observers represents the consistency of ratings by observers adjusted so that the differences between the means for observers is treated as irrelevant error. This method of calculating observer reliability is a better estimate in that it does not assign to error the tendency of observers to use the measurement scale differently, i.e., the tendency to rate higher or lower on the average. Since the scores from all observers were averaged to get an individual's score, these constant errors are "balanced out." The coefficient for the supervisors includes the constant errors as sources of unreliability.

The magnitude of the coefficients of reliability for the supervisors suggests that they were consistent in the judgments they made and that the eight supervisors used the scale in a similar manner. In spite of the difficulty in demonstrating a high reliability when data are "unmatched," the supervisors demonstrated a high consistency on all three scales. As would be expected the coefficients for observers were lower, for it would be more difficult to achieve a high "between judge" consistency among five observers rating the same subjects than to achieve a high average "within judge" consistency with each supervisor rating only his own subjects. Recognizing the differences in the two estimates of reliability, it was concluded that coefficients represented sufficient reliabilities to give meaning to further computations.

Table 6.2 also gives the product moment correlations between the ratings of supervisors and observers for the three variables. In each case the coefficient was significantly larger than could have been expected by chance. To obtain an estimate of the "true" correlation between the ratings of supervisors and observers, each correlation coefficient was corrected for the unreliability of both supervisors and observers. The corrected r's were found to be .82, .87, and .95, evidencing an extremely congruent ordering of individuals on the three variables. These results indicate that the greatest limitation on agreement between supervisors and observers was not their combined difference in background and method of observation but their low reliabilities.

Although it was concluded that there was a high relationship between the ratings of classroom performance by supervisors and observers, the investigators felt that some further comparative analyses of these data should be made. Since the ratings by supervisors of student teachers constitute a major evaluation of the teacher education graduate, it was decided that some investigation of the individual ratings should be made. Since the supervisors' ratings were "unmatched", i.e. only one supervisor rated a student teacher in a given quarter, there was an opportunity for bias to operate consistently for or against any one student. Since the observers' ratings were "matched" or balanced, a comparison was made between the supervisors' and observers' ratings of each student teacher. The observers' ratings were used as a basis of comparison without any thought that they should be interpreted as the "right" appraisal of a student teacher, for it might be more valid to argue that supervisors' ratings represented the "expert" judgment. The comparison was made because there was substantial general agreement between the two ratings and because the observers' ratings can be represented as the unbiased appraisals. To make the comparisons easier to interpret, both supervisors' and observers' ratings were adjusted by a linear transformation to a mean of 50 and a standard deviation of ten (Z scores). For the analysis, the mean of the observers' Z scores for each student teacher was subtracted from the mean of the supervisors' Z scores for the same student teacher to give a discrepancy score.

Table 6.3 presents the distribution of the discrepancy scores by supervisor for ratings of creativity. The data show that for seven of the thirty-six students the discrepancy was larger than one standard deviation of the Z scores, ten Z score points. Although no statement can be made as to whether the discrepancies represent errors, it would seem that there is a reason for concern. To determine whether the discrepancy scores were larger on the average for some supervisors than others, the analysis of variance test presented in Table 6.3 was made. The F value was not significantly larger than could be expected by chance indicating that there were no significant differences among the mean discrepancy scores of the eight supervisors.

Table 6.4 presents comparable data for the ratings of interpersonal relations. Seven of the thirty-six student teachers had discrepancy scores larger than ten Z score points. The analysis of variance test indicated that the discrepancy scores were not different among the supervisors.

Table 6.5 presents the analysis of the discrepancies in rating the success as a student teacher. Eight of the thirty-six discrepancy scores were larger than ten Z score points. The analysis of variance test indicated that there were significant differences among the mean discrepancy scores for the supervisors. Supervisors 5 and 8 tended to be considerably less different from the observers in their ratings than did other supervisors.

Table 6.3

Distribution by Supervisor and Analysis of Variance of Z Score Discrepancies (Supervisor's Mean Minus Counselors' Mean) for Creativity Scores

	Supervisor							
	1	2	3	4	5	6	7	8
Scores	-14	-27	23	13	-5	-6	-4	-1
	12	-9	8	2	3	-5	13	4
	-11	-7		-3	-1	7	11	1
		-6		8	2	-4	4	-6
		4		-18	-4	10	8	-11
				5				
Mean	12.3	10.6	15.5	8.2	3.0	6.4	8.0	4.6

Analysis of Variance

SV	df	SS	MS	F
Between Supervisors	7	396.622	56.660	1.92 n.s.
Within Supervisors	28	827.601	29.557	

Table 6.4

Distribution by Supervisor and Analysis of Variance of Z Score Discrepancies (Supervisor's Mean Minus Counselors' Mean) for Interpersonal Relations Scores

	Supervisor							
	1	2	3	4	5	6	7	8
Scores	-8	6	11	8	-2	-2	-9	2
	11	-8	-9	12	0	0	20	0
	-10	-5		-3	-8	10	11	-2
		-9		9	0	-6	3	-2
		-1		-4	-8	1	0	-12
				1				
Mean	9.7	5.8	10.0	6.2	3.6	3.8	8.6	3.6

Analysis of Variance

SV	df	SS	MS	F
Between Supervisors	7	188.049	26.864	1.252 n.s.
Within Supervisors	28	600.701	21.453	

Table 6.5

Distribution by Supervisor and Analysis of Variance of Z Score Discrepancies (Supervisor's Mean Minus Counselors' Mean) for Success as Student Teacher Ratings

	Supervisor							
	1	2	3	4	5	6	7	8
Scores	-4	1	11	7	3	-7	-5	-1
	14	-10	10	7	2	2	-6	-2
	-12	-4		-7	3	9	-10	2
		-3		21	-1	2	9	0
		2		-6	2	-11	0	5
				-8				
Mean	10.0	4.0	10.5	9.3	2.2	6.2	6.0	2.0

Analysis of Variance

SV	df	SS	MS	F
Between Supervisors	7	326.205	46.600	3.125*
Within Supervisors	28	417.434	14.908	

The analysis of ratings of classroom performance leads to several conclusions. First, observers and supervisors achieved a satisfactory degree of reliability (consistency) in making their judgments. Second, there were some differences in the way in which observers and supervisors used the rating scale; observers rated students lower, on the average, with less variability. Considering the entire group there were substantial correlations (agreements) between observers and supervisors, and when the coefficients were corrected for the unreliability of both raters the agreements indicated were high in all three areas rated. Third, in spite of achieving satisfactory reliability and substantial agreement, there were significant discrepancies in the ratings given to individual students. The mean of observers' ratings of individuals were compared with the mean of the supervisors' judgments, and on each scale about one-fifth of the students in the group had ratings that differed as much as ten Z score points, one standard deviation on the scale used to express the ratings. These wide discrepancies between the "balanced" ratings of observers and the "unbalanced" judgments of supervisors suggest the need to study the patterns of supervision of student teachers. Even though there was general agreement between observers and supervisors there were significant deviations that might have rather serious consequences for individual students. Although one may give the ratings of supervisors the heaviest weight because they were the experts in the field, the fact remains that supervisor's judgments were "biased" in that they were not "balanced" by the ratings of other supervisors. The one rating scale on which discrepancy scores were significantly greater for some observers than for others gives support to the concern for the inevitable differences among supervisors.

Fourth, data from this study gave emphasis to the need for further study of the methods of observation used in supervision and the background of experiences essential for supervisors. It seems significant that sensitive observers using a structured observational technique were in general agreement with the expert in elementary education. Further studies need to separate the observer and the method of observation to study the contribution of each.

Chapter VII

Summary and Conclusions

The Subject of the Study

For more than a decade, the College of Education has collected data about students seeking admission. Throughout this period the data included interview information and test scores as well as the academic record and speech and health ratings. Since this information has been used in making decisions about the admission of the students, it is reasonable to question the reliability and the predictive validity of the facts and judgments placed in the students' records as part of the screening procedure. An earlier study checked the reliability of the interview judgments. (Edson and Wilk, 1958). The present study looks at the validity question, relating the admissions data to the subsequent behavior of students at the time they are student teachers.

The study focuses on three general questions, the first of which may be stated as follows:

> What is the relationship between admissions data collected prior to the first professional course in teacher preparation and the subsequent observed behaviors of student teachers.

In order that the first question might be answered, it was necessary to control or estimate variables related to student teaching. The second question, growing out of this necessity, is stated as follows:

> What is the relationship of student teaching preference, and experience, and of the classroom situation to observed behavior of student teachers.

Since the focus of the study is on the student teachers and the data that had earlier been collected about them, and since the concern of the study was with the relationship of the individual differences among student teachers as these related to their behavior, it was important that other factors affecting the assessment of classroom behavior be separated out. Methods of observation and choice of observers therefore required study. This led to the third general question which is stated in two parts as follows:

(1) Using the personality variables concerned in this study as the basis for comparison, what is the relationship of ratings of student teachers made by experienced student teacher supervisors using their established observational methods and the ratings derived from the observational techniques designed for this study and employed by research observers.

(2) Using certain teaching behaviors as the basis for comparison, what is the relationship of ratings of student teachers by student teaching supervisors using their established observational methods and the ratings made by research observers using the methods designed for this study.

The interview reliability study, to which reference has been made, made use of a view of personality that employed needs constructs in formulating hypotheses concerning individual behavior. Two propositions were central to this earlier study: first, that persistence in teacher education required at least two qualities, Achievement potential and need Achievement; and second, that interest in the field of teaching is related to the variables of need Dominance, need Integration, and Security. Two of these, need Dominance (n Dom) and need Integration (n Intg), are important to the present study.

At the time of the interview study, the counselors made ratings of each student on these personality characteristics. The behavior associated with n Dom and n Intg is presumed to be similar to behaviors that Anderson (1939) has described and that which Flanders (1960) has specified as "direct influence" and "indirect influence." Teacher behavior is said to have "direct influence" when the effect is to restrict the pupils' freedom of action. Such behavior is considered to be evidence of a high degree of n Dom. Teacher behavior is said to have "indirect influence" when the teacher encourages the pupil to expand on his own ideas or to determine the way in which he might respond. Teacher behavior that permits more freedom of action on the part of the pupil reflects a high degree of n Intg. It is presumed that teachers with different degrees of need in these two areas will, in general, evidence differences in their verbal behavior as they attempt to influence pupil behavior. Thus the verbal behavior of student teachers became the criterion for testing the validity of counselor judgments.

The interview reliability study was based on interviews conducted with women students applying for admission to the first professional course sequence of the elementary education program, which begins in the junior year. Information collected about these students provided data for the current study.

In the senior year elementary education students were placed in Minneapolis and St. Paul, Minnesota schools for student teaching during morning hours of two successive quarters. Criterion observations in the classroom against which interview data were judged, were made in the senior year with these women in elementary education as the subjects.

The Design of the Study

The Sources of Variability

Five sources of variability that could affect observational data were identified in the study. Some of these sources were treated in such a way that the amount of their effect on the observations could be estimated; others were treated only in such a way as to insure that they would not bias the results. The five sources of variability are: the student teacher, the pupils in the classrooms, the placement of student teachers in the classrooms, the admissions data, and the observations by supervisors.

The fact that the student teachers were of the same sex and of a similar academic background reduced two important sources of variation, but the fact that the students were dissimilar in their preference for teaching at the upper elementary level or the lower elementary level and the fact that students would be observed in the first and second quarters of student teaching, that is, after different amounts of experience in the classroom, could contribute to variations. This led to the decision that grade level preference and experience should be incorporated in the design.

The pupils and the classrooms considered together as a unit were another source of variation. Therefore, the design took into account the grade level and the location of the schools. The design was based on the assumption that observed behaviors would vary depending upon the grade level of the classrooms and the socio-economic characteristics of the neighborhoods in which different schools were located. The upper three grades and the lower three grades formed two gross subdivisions within the elementary school on which estimates of the effect of grade level might be made. Because the socio-economic level could not be incorporated into the design so that the extent of its' effect could be identified, this factor was controlled at the time of assignment of student teachers to insure that it did not bias the observations.

Minneapolis and St. Paul were divided into four geographical areas and adjustments were made between these to insure against socio-economic bias. When the boundaries were established on the basis of census information, two statistical tests were made to determine whether each of the areas was generally similar to the others, and to see whether the variability was similar within each. The number of student teachers living in each area or having ready access to it determined the number of schools within the area to be included in the study. Next, the student teachers were placed so that each had an equal opportunity to be assigned to any one of the classrooms in her area. This randomization assured that chance alone, not bias, could cause the socio-economic variable, as well as others not specified, to be differentially effective on classroom observations.

The admissions data employed in the study were of two types. First, there were the psychological tests which include the *Miller Analogies Test*, the *Co-operative Reading Test*, Form C2R, the *Minnesota Teacher Aptitude Inventory*, the *Minnesota Multiphasic Personality Inventory*, and the *Teacher Opinion Inventory*. Second was the set of counselors' judgments. These judgments were based on test scores and interview information combined.

The study utilized two sets of judgments about student teacher behavior: first, there were those of the student teaching supervisors regularly employed by the College of Education in the student teaching program and utilizing the observational procedures regularly employed by the office of student teaching; second, there were the observations of the observers employed for the purpose of this study using the observational procedures set up especially for the purpose of this study. The supervisors and the observers made judgments in the same time period. Their judgments were used as a basis for comparing these methods of observation.

The Cross-over Design

The design treated four sources of variation as experimental variables. These were 1) student teachers' grade level preference, 2) student teachers' quarter of experience, 3) classroom grade level, and 4) order of placement in two different grade level assignments. The controlled variables, that is, those treated in such a way that they would not bias the observations, were: 1) the professional education of the student teacher, 2) the sex of the student teacher, 3) travel by the student teacher, and 4) socio-economic level of the classroom. A diagramatic representation of this cross-over design employed in the study is given in Figure 2.1. This cross-over design showed that the effect of each of the four placement variables could be isolated, and that those factors that were controlled would not bias the observation.

Observation and Measurement of Variables

The predictor variables employed in the study have been mentioned. There remain for discussion the observed variables and the concurrent variables. The observed variables are of two kinds: those describing classroom activities and those classifying verbal behaviors.

Unlike the approach to the description of classroom activities that is frequently employed, this study did not call for the ratings or the judgments of supervisors in dealing with the classroom activity. Rather, it employed a modified form of the Mitzel and Medley observation schedule and record, the OScAR III. This observation schedule and the manual for its use are in Appendix A. As they used it, observers checked whether each behavior occurred in each five-minute observation period. In a sense, each observation category or item is analagous to a test item which is answered "yes" or "no" for each five-minute observation period.

To collect data concerning classroom verbal behaviors, the observers employed a form of the Minnesota System of Interaction Analysis developed by Flanders. Modifications of this system as used in this study are described in Chapter 2. The system employs four categories of direct influence and four categories of indirect influence that describe the teachers' communication to the class concerning primary learning tasks. Direct influence and indirect influence are assumed in this study to be evidence of behavior related to n Dom and n Ing respectively. In each visit to a teacher, the observers recorded the verbal behavior over three separate five-minute observation periods. The variables of concern, n Dom and n Intg, were represented for a five-minute period as the sum of the frequency per hundred for the four relevant categories. The score representing the visit of the observers was the sum of the scores obtained for the separate five-minute periods. The categories were weighted differentially according to the contribution of each category to the total score by means of reciprocal averages. The weighting system is discussed in greater detail in Chapter 3.

The third of the three questions that were the subjects of this study concerned the relationship between the ratings made by the experienced student teaching supervisors using their methods of observation, and the data obtained by the research observers using the methods selected for this study. The judgments of behavior made by the student teaching supervisor were the concurrent variables of the study. It should be noted that in seeking an answer to the third question, one cannot extract the separate effects of the experience and backgrounds of the raters from the methods of observation employed, for these two factors are inseparable in this study.

For the supervising teachers, the variables were determined by their ratings on the five-point scale that appears in Appendix A. After the supervisors had concluded their observations as they would normally conduct them, they completed a rating on the n Dom and n Intg scales. Each supervisor observed the same student at least three times in the course of a quarter. Comparisons were based on the composite judgments. For observers, the variables were the n Dom and n Intg scores derived as described above.

To answer the third question in its second form required that an instrument having relevance for the objectives of student teaching be constructed. A copy of this instrument, "Ratings of Classroom Behavior," is included in Appendix A. It contains two sections: the first, headed "Creativity and Resourcefulness in Teaching," and the second, headed "Interpersonal Relations." Both the supervisors and the observers completed this twenty-one item scale. In addition, the supervisors and observers made an estimate of the success in student teaching and an estimate of the probable success in teaching, indicating these on a five-point scale.

The Conduct of the Study

In preparation for this study, the directors consulted with the appropriate school officials and met with teachers to enlist their cooperation. Following this, they met with student teachers who were available for placement to explain to them the purposes of the reseach project. The student teaching supervisors attended the meeting with the prospective student teachers.

During the four week period prior to the first observations, the observers, who had been selected on the basis of their background in psychology and the behavioral sciences, as well as on evidence of their sensitivity to inter-personal relations, were employed on a half-time basis in order that they might be trained in the techniques of observation. After the observers began their observations of the student teachers, they made periodic reliability checks by going as a group to elementary classrooms, as they had done during the training sessions.

The plan for the randomized assignment of student teachers within the four geographical areas has already been presented, briefly in this chapter, and in detail in Chapter 2. Thirty-eight student teachers, selected from the 167 students who participated in the interview study, were assigned according to the plan. Of the 38 student teachers, two were dropped from the sample before the end of the study. It was also necessary to replace two supervising teachers. In making these adjustments, the investigators were careful to protect the requirements of the design.

Each of the five observers visited each student teacher once during each quarter. Thus there were ten different observation reports for each student teacher during the two quarters in which the study was conducted. The plan called for an average of 8.3 observations per day with three to four days between two observations in the same classroom during the fall quarter. The plan for the winter quarter called for approximately 5 observations per day, each student teacher to be visited once a week. In practice, the time between observations during the fall quarter ranged from zero to 14 days, and during the winter quarter from zero to 12 days. In no case, however, were two observations made simultaneously.

The Observed Behaviors of Student Teachers

Two instruments were employed in the observation of student teacher behavior: the OScAR III and the Minnesota System of Interaction Analysis. Although these instruments are presented in detail in Chapter 3, they are summarized here.

The OScAR III

The individual items of the OScAR III were combined to form eight scales. For purposes of this study, the scale scores for every observation for all student teachers were computed on the Univac Scientific Computer (ERA 1103). In order that individual scores on the various scales might be compared, a linear transformation was performed to set the mean scale value at 50 and the standard deviation at 5. The scale titled Classroom Variety of Materials (CVM) is composed of items about two kinds of materials that were used by the teacher or the pupils: the verbal materials, such as text books or workbooks; and the audio-visual materials, such as blackboards, maps, films, or records. The items deal with whether the materials were used rather than the percent of the time they were used. Consequently, the total score for this scale, as for all scales of the OScAR III, is an indication of whether the activity was observed during the visit to the classroom.

Another scale, Teacher Class Activities (TCA) contained a series of items that were checked when the student teachers took an active lead in directing the class and when pupils were, more or less, passive participants. Teacher Grouping Activities (TGA) is an indication of the extent to which the student teacher grouped pupils. The scale titled Teacher-Individual Activities (TIA) attempts to describe the activities of the teacher that are directed toward individual pupils to encourage them individually in their learning efforts. Another group of items forming the Pupil-Class Activities scale (PCA) is a record of the events in the classroom in which individual pupils who are reciting, reporting, or leading the class are the focus of class attention.

One of the scales deals with pupil response to the teacher-initiated activities recorded in other scales. This scale, titled Pupil Disruptive Behavior (PDB) is an indication of whether pupils are inattentive, reluctant to cooperate, or resistant to student teacher efforts. Two additional scales, Classroom Emotional Climate (CEC) and Classroom Verbal Emphasis (CVE) contain a large percentage of item overlap with the scales that have been listed. These scales were suggested by the factor analysis of classroom observation data made by Medley and Mitzel.

Intercorrelations among the five scales first mentioned were not significantly larger than could be expected by chance. It may be concluded that they measure different kinds of classroom activities. These five scales were combined to form the Classroom Activity Index (CAI) which is a reflection of the variety of activities observed. When each of these five individual scales and the composite scale were correlated with Pupil Disruptive Behavior (PDB) all of the correlations were negative and two were significantly larger than would be expected by chance. In general it may be said that the wider the variety of activities observed, the less disruptive were the pupils. The strongest relationship was between TIA and PDB. T e

-.63 correlation suggests that the more frequently the student teacher was observed to be interested, helpful and supportive, the less frequently the students were observed to be disruptive.

The Minnesota System of Interaction Analysis

The form of the Minnesota System of Interaction Analysis used employs nine categories into which teacher verbal behavior may be classified. Four of these are categories of indirect influence, as the acceptance of feeling or use of praise and encouragement, and four are categories of direct influence, as the use of criticism or the justification of authority. A ninth category of teacher talk relates to administrative or routine procedures which are not related to the primary learning tasks of the class.

In order to draw any conclusions based on observations employing the instrument, it was necessary to determine its reliability and to make a comparison of weighted and unweighted scoring methods. Because observers went to classrooms individually to observe the subjects of the study, it was necessary to determine the reliability of their observations apart from those they conducted to answer the principle questions of the investigation. For this reason the reliability estimates are based on observations made when the observers went as a group to classrooms during the training sessions prior to the observations made for the study, and, when they went together to classrooms at intervals during the study to conduct reliability checks. Determination of the reliability estimates were made from an analysis of variance model which specified the sources of variation that could affect the scores resulting from the classification of verbal behaviors. The reliabilities of the pooled ratings of the five observers based on scores of n Intg ranged from .90 to .97 with a mean of .94. For n Dom scores, the values ranged from .90 to .98 with a mean of .95. It was inferred that the score values computed for each subject had sufficient reliability to give stable estimates for the variables observed.

Since the categories of the Minnesota System data were combined, one could reason that certain categories might be more potent contributors to the defined concepts of n Dom and n Intg than others. One could also argue that these categories which were thought to be most potent might be counterbalanced by the more frequent occurrence of the other categories. It was necessary, therefore, to study the appropriateness of weighted and unweighted scores. The technique of reciprocal averages was used to derive the weighted scores. Analysis of the data is discussed Chapter 3. A comparison of the a priori and the reciprocal average weights supported the assumption of a differential potentcy of the verbal behavior categories.

The Relation of Observed Behaviors to Classroom Placement Variables

The basic question in this study concerned the validity of admissions data as predictors of student teacher behavior, but before such a question could be answered, another question had to be posed. This prior question asked whether there was a relationship between the observed behaviors and such classroom placement variables as teacher preference for upper or lower grades, teacher experience, or classroom characteristics. The question concerned the data obtained by means of the OScAR III and the Minnesota System, and utilized a cross-over design to test for the effect of the sources of variability identified. The answer to this question was required to indicate whether the criterion variable behaviors were affected by situational variables.

It will be recalled that the experimental design treated four sources of variation associated with classroom placement: student teachers' grade level preference, student teachers' quarter of experience, classroom and grade level, and order of placement in two different grade level assignments. Subjects eligible for placement in the experiment were first divided into two groups according to their stated grade-level preference, upper grades or lower grades. Each of these two preference groups was then randomly assigned to one of the two possible orders of experience, and one-half of each preference group received its stated choice the first quarter. The second quarter of student teaching, the level at which the teaching was done was reversed. The analysis of variance tables for this cross-over design are presented in Chapter 4.

Analysis of the OScAR III Scales

Results of the analysis of the TIA scores indicate that when in the upper grades, student teachers were more often observed giving positive attention and encouragement to individual pupils than when they were in the lower grades. At the upper level teachers more often used direct questions to a pupil or answered individual pupil's questions rather than ignoring them.

Data from the TCA scale led to the conclusion that as student teachers gained experience, they took more active lead in directing the class, and students became more passive participants. Teachers more frequently lectured, told stories, read to the class, demonstrated, led singing or games, and illustrated at the board. The increase in these activities was greater for that group of teachers which moved from the upper grades to the lower grades in the second quarter of experience.

In one section of the schedule, PCA, observers recorded the events in the classroom in which individual pupils were the focus of class attention. The activities were similar in nature to those in which the student teacher led the class, except that the leadership came from the pupils rather than from the student teacher. The variety of activities in which pupils led the class were larger when student teachers were in the lower grades than when they were in the upper grades.

From the analysis of the CVM scale, one may conclude with regard to the use of verbal and audio-visual materials that student teachers and pupils made more use of a wider variety of materials when they were in the upper grades than when they were in the lower grades. Student teachers were more frequently observed using audio-visual materials than verbal materials, but pupils more frequently used verbal materials.

Two OScAR III scales remain: one dealing with the formation of groups of students by the student teachers, TGA, and one dealing with pupil disruptive behavior, PDB. None of the experimental variables was found to have a significant effect on these scales. From the scales that have been discussed, two composite scales were prepared as suggested by Medley and Mitzel. One of these dealt with the classroom emotional climate. The analysis indicates that those students who preferred upper grades had significantly better Classroom Emotional Climate (CEC) scores when they were in the upper grades, but that those who preferred the lower grades showed no change in scores at the two different levels.

Those who preferred the upper grades and taught in the upper grades during their first quarter of student teaching achieved a less favorable emotional climate in the lower grades (their second quarter) than their counterparts who taught first in the lower grades.

The other composite scale dealt with the classroom verbal emphasis. Those student teachers who preferred the lower grades had a higher mean Classroom Verbal Emphasis (CVE) score than those who preferred the upper grades, but it may be said that all student teachers, regardless of preference, had higher verbal emphasis scores when they were in the upper grades.

Since the student teachers' preference, teaching level, and experience were not related to the disruptive behaviors of pupils, there remained the possibility that pupil inattention, reluctant cooperation, or resistence to the teachers' effort might be associated with classroom activities. Therefore, the disruptive activities were studied in the context of teacher-class activities, teacher-grouping activities, teacher-individual activities, pupil-class activities, and variety of materials used by teachers. Taken individually, two of the activities are related to disruptive behavior beyond the level expected by chance. In those classrooms where the level and variety of pupil-class activities were high, pupil disruptive behavior tended to be lower, and in those classes where the student teachers' work with the individuals was observed to be interested, helpful and supportive, disruptive behaviors were fewer. Other activity variables were not related to disruptive behavior beyond the level expected by chance, but all were related in the negative direction, indicating that the more frequently the various activities were observed, the less frequently pupil disruptive behavior occurred.

Analysis of Verbal Behavior

Since n Intg and n Dom scores served as criterion variables, it was essential to find whether these variables were affected by the placement variables. If the situation in which some groups of student teachers were placed caused them to respond differently than other groups then some adjustment for the situational variable would be required. In terms of the design these differences would be found in groups who had different grade level preference or order of experience. If, however, the situation caused all students to change, from one quarter to the next, for example, then no adjustment would be needed.

In Chapter 4, the analysis indicated no relationships between the placement factors and the n Intg scores when raw scale scores were used. When the reciprocal averages technique was employed, the scores were significantly higher the second quarter of student teaching.

With respect to n Dom, the analysis of the raw scores and the reciprocal averages scores revealed three significant differences; however, an inspection of the means concerned suggested that the differences could not be given a straightforward interpretation, but resulted from an interaction of level and order effects. The two significant main effects and the interaction was caused by the mean score of one group, namely, the group assigned the lower-upper student teaching order as they engaged in the first quarter of student teaching (at the lower level). Since the one mean alpne accounted for all the differences observed in the analysis, one may question what account should be taken of this in further analyses that employed n Dom as a criterion variable. The question is raised and discussed, and the conclusion is given in Chapter 4. For purposes of this study, it was assumed that the differences found were not straightforward "between groups" differences, and so in further analyses, no adjustments in n Dom scores were made.

The Comparison of Observed Behaviors and Ratings by Supervisors

One can suppose that the relationship between predictions made at the time students are admitted to the first education course and the assessments made in the senior year when the student is enrolled in student teaching might be influenced not only by the student teaching placement variables, which have been discussed, but also by the methods of observation employed during student teaching as well. The fact that the study employed student teaching supervisors, following methods similar to those generally used in student teaching in many college programs, and trained observers following methods selected specifically for this study made it possible to examine some questions concerning methodology.

The third general question of concern in this study, which was stated in full at the beginning of this chapter, asks the relationship between the ratings of the student teaching supervisors and the research observers when their attention is directed first to personality variables and second to teaching behaviors. If no relationship were found, this might be attributed to a lack of definition of the variable as well as to differences in the observation and measurement methodology, but if a significant relationship were found, it might be assumed that the variables had sufficient definition to be detected by the two different groups of observers using two different observational and measurement procedures.

The student teaching supervisors and the research observers were given definitional statements of the personality variables, n Dom and n Intg, at the beginning of the study. The supervisors' judgments of these variables were quantified by means of a rating scale. Procedures employed by the research observers have been described. In the analysis both the unweighted and the weighted scores obtained from the observers were examined.

In general, there seems to be little relationship between the n Dom scores which represented the observations of the two groups, supervisors and observers. Although the correlation was significant when the winter-quarter, weighted, observers' scores were examined, the relationship was not a significant one when both quarters were considered together.

In the case of the second personality variable, n Intg, the outcome was different. In general it can be said that the agreement between the scores of supervisors and observers was significant, for it was greater than chance expectations.

In order that the agreement concerning teacher behaviors might be examined, the supervisors and observers used a measuring scale constructed especially for this study. The purpose of the instrument was to assess the extent to which two objectives of student teaching were observed: 1) the development of creativity and 2) the achievement of harmonious interpersonal relationships. The supervisors and observers also rated overall effectiveness in student teaching and predicted success in teaching. The reliability of the supervisor and observer ratings were first examined. The design of the study made it necessary to examine the consistency with which each supervisor rated those students he alone supervised. This gave a within-judge measure of consistency. Since the research observers saw all students, the reliability of their ratings was based upon between-judge consistency. The magnitude of the coefficients of reliability for the supervisors suggests that they are consistent in judgments they make and that the eight supervisors used the scale in a similar manner. As might be expected, the between-judge consistency for the research observers was slightly lower. Recognizing the differences in the two estimates of reliability, it was concluded that the coefficients represented sufficient reliability to give meaning to further computations.

The agreement between supervisors and observers were then examined. In each case the product moment correlation coefficient was significantly larger than could have been expected by chance. When the coefficients were then corrected for the unreliability of both supervisors and observers, the coefficients were found to be .82, .87, and .89, which may be interpreted to mean that the greatest limitation on agreement between supervisors and observers was low reliability rather than differences in background and method of observation.

Since the ratings of supervisors of student teachers constitute a major evaluation of the graduate, and since only one supervisor rates a student teacher in a given assignment, there is an opportunity for bias to operate consistently for or against any one student. For this reason the investigators felt that some further comparative analysis examining individual ratings should be made.

The composite score obtained from research observers and from the individual ratings of the supervisors were converted to standard Z scores. The discrepancy between the observer and the supervisor scores for each student was the basis for the analysis.

The data show that for seven of the thirty-six students the discrepancy score for the ratings of creativity was greater than ten Z score points, one standard deviation of the scores. On the basis of comparable data for the ratings of interpersonal relations, seven of the thirty-six student teachers had discrepancy scores larger than ten Z score points. Further analysis indicated that the discrepancy scores were not different among the supervisors.

The Relation of Predictor Variables to Observed Behaviors

The basic question from which this project started dealt with the relationship between admissions data collected prior to the first professional course in teacher preparation and the subsequent observed behaviors of student teachers. The admissions data to which the question refers includes high school rank, honor point ratio, Miller Analogies Test, Minnesota Teacher Attitude Inventory, (Bowers') Teacher Opinion Inventory, and the counselors' judgments concerning n Dom and n Intg. The observed behaviors, i.e. the criterion variables, are the verbal behaviors observed by means of the Minnesota System of Interaction Analysis that are assumed to be evidence of n Intg and n Dom.

The study reports the relationship among the predictor variables as well as between these variables and the criterion scores. Since the most probable application of the results of these analyses would be to predict teacher behavior from admissions data, the study employs the multiple regression procedure to suggest whether the variables are predictive when used in a linear combination.

Of the thirty-two correlations among predictor variables and criterion variables, only five are larger than would be expected by chance. High school rank and the Miller Analogies Test correlated significantly with sophomore honor point ratio, and the Miller Analogies Test correlates significantly with the Teacher Attitude Inventory. Honor point ratio and the Teacher Attitude Inventory scores were related to the n Intg score beyond chance expectations.

The table of intercorrelations suggested that a linear combination of predictors might be successful in predicting the criterion variables. Following appropriate analyses as reported in Chapter 5, two multiple regression equations were computed, one for each of the two different observed teacher behaviors. The first regression equation was computed using n Dom (RA) as the predicted variable. The attained value for R was .3077. An F test was computed to test the hypothesis that the combined relationship of the predictor to the predicted variables could be accounted for by chance, and this hypothesis was accepted at the .05 level of significance. The admissions data were not significantly related to classroom observations of n Dom.

In the second regression equation n Intg (RA) was the predicted variable. The obtained value of R was .5987. The hypothesis that the relationship could be accounted for by chance was rejected at the .05 level of significance. It may be concluded that a significant linear relationship between the combination of independent variables and the dependent variable n Intg, does exist. When the individual contributions of each predictor variable were studied and non-contributing variables were deleted, the sophomore honor point ratio and the Minnesota Teacher Attitude Inventory score remained as the two variables which accounted for the relationship.

Conclusions

The study deals with three general questions that ask about (1) the validity of admissions data, (2) the effect of different student teaching placements, and (3) the relationship between the ratings of student teachers made by student teaching supervisors and the ratings made by research observers. In seeking answers to the main questions, research workers made observations that enable one to comment on other topics which are secondary to the three central issues. Among the topics of secondary interest in this study--but of significance for student teaching supervision--are these: the nature of student teacher activities, the materials used in the classroom, the pupil response to student teachers, the reliability of student teaching supervisors' ratings, and the verbal behavior of student teachers.

Conclusions about the general questions are presented first. The principal question asked about the validity of admissions data as predictors of the n Intg and n Dom observed in the behavior of student teachers. The study gives support to predicting n Intg from certain items of admission data, namely Minnesota Teacher Attitude Inventory scores and sophomore honor point ratio. In explaining this, one may speculate that the students who had greater command of the subject matter they were teaching and who had certain attitudinal dispositions were comfortable enough in the classroom to permit the students more freedom to respond. They could accept pupil feeling and they could support, praise and encourage pupils.

It is equally important to note that counselors' estimates were not significantly related to student teaching performance. On the basis of the admissions interview and data available at the time of admission, counselors could not predict successfully which students would exhibit integrative classroom behavior as student teachers.

The construct of n Intg as observed in this study, is unrelated to the placement variables and is sufficiently well defined to be observed consistently and with high agreement by observers and supervisors. These findings indicate that n Intg has the stability required of a criterion and in addition, it has relevance to the problem of judging the validity of admissions data.

The construct, n Dom, was not found to be such a stable criterion. Although as defined it has relevance to the problem studied, n Dom was found to be related to some placement variables, and the low degree of agreement between observers and supervisors suggests that the concept did not have the same meaning for these two groups.

Classroom placements were found to be related to some types of student teaching performance. The classroom placements appear to be more closely related to specific student teacher responses to individual pupils and to the class than they are to any generalized tendency of student teachers to use certain categories of verbal behavior. For example, in the lower grades the student teachers were more likely to exert direct influence on pupils (to be critical or to give directions), and in the upper grades, to exert indirect influence (to be supportive, and helpful and to accept the feelings of students).

The student teaching supervisors and the research observers, that is, the experts and the trained observers, exhibited a high degree of general agreement when rating either personality variables or judging classroom performance. One may conclude from this that the skilled research observers are not pedagogically naive, and that individuals such as these, given appropriate supervision and direction, may be employed to assess the behavior and the activities of student teachers.

To the extent that there is a lack of agreement between the ratings of supervisors and observers one may conclude that this is due to the unaccounted-for variance, i.e., error, but one may also interpret the data to suggest that rater bias is an important factor in the deviations when there is but one rater to judge a student's performance, as occurred in student teaching supervisor ratings. The skilled research observers, who were not presumed to be experts in the supervision of student teachers, were as reliable in their ratings as the regularly employed student teaching supervisors and, in general, they agreed with the supervisors in their ratings.

The conclusions stated up to this point have dealt with the three principal questions of the study. The following comments concern topics of secondary interest relative to the purposes of the study, but of general importance in the supervision of student teaching: the frequency and effect of specific kinds of classroom activities and behavior.

Student teachers, aware of difficulties that may arise, may be quite apprehensive about discipline as they prepare to meet their first classes. Within the limits of the current study, one can make some statements about pupil cooperation as student teachers learn the art of teaching. One conclusion is that the frequently observed disruptive behavior was an annoyance rather than a threat to student teacher control. Whispering, laughing, or interrupting occurred frequently, but hostile acts seldom. It may also be inferred that there is a clear relationship between method and pupil behavior. When student teachers employed a variety of procedures, disruptive behavior was less frequent. The more the student teacher was supportive, encouraging, and helpful, the fewer were the disruptive behaviors. The amount of disruptive behavior is sometimes presumed to be related to the age of the children, but there was no difference in the amount of disruptive behavior at the lower and the upper grade levels.

In view of the relationship between student teacher behavior and pupil response, it is worthy of note that although student teachers were observed to be praising good work and assisting individuals in over half of the visits, in eightly percent of the visits they were critical or sarcastic about pupil behavior not directly related to the learning situation.

Observers also made note of the methods used to direct learning. In over half of the visits, student teachers presented materials in a formal way, that is, outlining, telling pupils about the topic, or summarizing. They made formal presentations in the manner of a lecture more frequently in their second quarter of student teaching than in their first. Student teachers used a greater variety of materials and directed activities to the class and to individual pupils more frequently in the upper grades than in the lower grades. At the upper level the pupils more often recited, talked to the class, or read aloud.

In summary, as one examines the many aspects of the data to find a pattern, if such exists, it appears that the student teacher behavior said to define need Integration, which is somewhat predictable from sophomore honor point ratio and Minnesota Teacher Attitude Inventory scores, is also associated with fewer pupil disruptive acts and a learning climate characterized by a greater variety of teaching materials and more frequent pupil activity.

REFERENCES

1. Anderson, H. H. The measurement of domination and of socially integrative behavior in teachers' contacts with children. Child Develpm. 1939, 10, 73-89.

2. Anderson, H. H., Brewer, J. E. & Reed, M. F. Studies of teachers' classroom personalities III: Follow-up studies of the effects of dominative and integrative contacts on children's behavior. Psychol. Monogr. 1946, No. 11.

3. Anderson, H. H. & Brewer H. M. Studies of teachers' classroom personalities II: Effects of teachers' dominative and integrative contact on children's behavior. Psychol. Monogr. 1946, No. 8.

4. Baker, F. B. Simplified test analysis program. Research Memorandum, 1959, BER-59-18, Bureau of Educ. Res., University of Minnesota.

5. Cochran, W. S. & Cox, Gertrude M. Experimental designs. New York: John Wiley & Sons, 1950.

6. Cornell, F. G., Lindvall, C. M., & Saupe, J. L. An exploratory measurement of individualities of schools and classrooms. Bureau of Educ. Res., College of Ed., U of Ill., 1952.

7. Edson, W. H. & Wilk, R. E. An experimental study of the admissions interview in teacher education for predicting success in teaching. (Multilith) College of Education, University of Minnesota, 1958.

8. Flanders, N. A. Teacher influence, pupil attitudes and achievement. (Multilith) College of Education, University of Minnesota, 1960.

9. Lowenfeld, Viktor. Current research on creativity. NEA Journal, 1958, 47, 538-40.

10. Medley, D. M. & Mitzel, H. E. A technique for measuring classroom behavior. J. educ. Psychol., 1958, 49, 86-92.

11. Medley, D. M. & Mitzel, H. E. Studies of teacher behavior: the refinement of two techniques for observing teachers' classroom behaviors. Res. Series No. 28, Division of Teacher Education, Board of Higher Education of the City of New York, 1955.

12. Mitzel, H. E. & Hoyt, C. J. A methodological study of reciprocal averages technique applied to an attitude scale. J. counsel. Psychol., 1954, 1, 256-259.

13. Mitzel, H. E. & Rabinowitz, W. Assessing social-emotional climate in the classroom by Withall's technique. Psychol. Monogr., 1953, 67, No. 18, (Whole No. 368).

14. Mosier, C. J. Measurement in rural housing, a preliminary report. Educ. psychol. Measmt. 1942, 2, 139-152.

REFERENCES con't.

15. Murray, H. A. *Explorations in Personality*, New York: Oxford U. Press, 1938.

16. Smith, B. O. A concept of teaching. *Teach. Coll. Rec.*, 1960, *61*, 229-241.

17. Stanley, J. C. Statistical analysis of scores from counterbalanced tests. *J. exp. Educ.*, 1955, *23*, 187-207.

18. Walker, Helen M. & Lev, J.. *Statistical inference*, New York: Holt, 1953.

19. Withall, J. G. The development of a measurement of social-emotional climate in classrooms. *J. exp. Educ.*, 1949, *17*, 347-361.

20. Withall, J. G. The development of the climate index. *J. of educ. Res.*, 1951, *45*, 93-100.

APPENDIX A

A-1 - Socio-economic Data

A-2 - University Supervisors' Rating Forms

A-3 - The OScAR III

A-4 - The Minnesota System of Categorizing Verbal Behaviors

Appendix A-1

Table A-1.1

CENSUS TRACT, OR PARTS OF CENSUS TRACT, AREAS SERVED BY ELEMENTARY SCHOOLS USED IN STUDY

SCHOOL	CENSUS TRACTS SERVED
Minneapolis	
1.	23, 29, 34, 35, 42, 45
2.	102, 103, 111
3.	98, 106, 112, 113
4.	21*, 28, 29, 33, 34
5.	32, 33, 41
6.	72, 73, 74, 79, 85, 86
7.	118, 121
8.	113, 114, 120
9.	1
10.	74, 75, 76, 88, 89, 104
11.	2, 3, 4*, 7
12.	13, 14, 20, 21, 22
13.	53, 54, 58, 59, 69
14.	111, 119, 121
15.	4, 8, 9, 10
16.	49
17.	110, 111, 117
18.	49, 50
19.	5, 6, 17, 18, 24
20.	83, 94, 95, 100
21.	68, 69, 70, 71, 77, 78, 82, 83
22.	115, 116, 120
St. Paul	
23.	32, 33, 48, 49, 50
24.	66, 75
25.	58, 59, 60, 68, 69
26.	15, 30, 43
27.	53, 56, 57, 58, 66, 67, 68
28.	14, 30
29.	1, 19

*Not included in socio-economic computations due to the small census tract area (less than one residential block) served by the elementary school.

Table A- 1.2

Z SCORES OF MEDIAN EDUCATION, MEDIAN INCOME, MEDIAN RENT AND POPULATION DENSITY PER RESIDENTIAL PLUS COMMERCIAL ACRE* FOR EACH CENSUS TRACT USED IN MAKING SOCIO_ECONOMIC COMPUTATIONS FOR SCHOOL AREA SERVED BY ELEMENTARY SCHOOLS USED IN STUDY.**

CENSUS TRACT NUMBER	EDUCATION (Note a)	INCOME (Note b)	RENT (Note c)	POPULATION DENSITY (Note d)	TOTAL (Note e)
Minnespolis					
1	46.64	52.60	56.42	65.18	220.85
2	48.88	57.64	50.89	64.36	222.77
3	51.87	59.80	57.34	67.06	236.07
4	40.67	52.73	44.34	67.10	204.84
5	36.19	54.38	40.91	62.07	193.55
6	50.37	54.11	50.57	63.06	218.11
7	51.87	54.37	54.34	66.24	226.82
8	49.63	58.43	52.73	64.40	225.19
9	44.40	54.52	48.41	65.04	212.37
10	35.45	52.74	39.38	62.94	190.51
13	37.69	54.46	48.56	63.89	204.60
14	46.64	59.00	48.29	65.04	218.97
17	35.45	52.38	38.62	63.01	189.46
18	36.94	51.52	41.32	62.27	192.05
20	46.64	53.39	50.89	63.78	214.70
21	44.40	51.74	42.79	62.48	201.41
22	46.64	48.75	42.73	61.31	199.43
23	34.70	46.12	36.87	57.35	175.04
24	33.96	49.23	37.21	60.25	180.65
28	55.60	50.29	48.54	62.67	217.10
29	40.67	45.87	43.12	58.34	188.00
32	50.37	59.76	50.54	63.24	223.91
33	42.16	48.01	50.69	59.14	200.00
34	36.19	41.37	42.57	42.51	162.64
35	33.96	41.08	34.03	58.70	167.77
41	35.45	50.00	42.83	58.09	186.37
42	33.21	38.97	37.36	54.08	163.62
45	30.22	23.20	37.24	61.03	151.69
49	72.01	26.10	56.39	66.60	221.10
50	66.79	54.31	55.32	67.53	243.95
53	54.85	35.35	45.54	39.12	174.86
54	36.19	32.00	42.93	32.23	143.35
58	42.16	39.37	42.93	27.81	152.27
59	44.40	37.14	43.29	44.55	169.38
68	61.57	42.72	52.87	53.09	210.25
69	50.37	39.88	44.52	54.72	189.49
70	59.33	45.15	51.34	57.99	213.81
71	56.34	42.41	47.85	52.12	198.72
72	39.93	46.80	42.89	57.46	187.08
73	39.93	48.21	43.13	59.10	190.37

Table A-1.2 con't.

CENSUS TRACT NUMBER	EDUCATION (Note a)	INCOME (Note b)	RENT (Note c)	POPULATION DENSITY (Note d)	TOTAL (Note e)
74	34.70	47.56	39.29	62.78	184.33
75	37.69	52.13	46.34	65.40	201.56
76	54.10	59.02	51.21	67.97	232.30
77	48.13	47.12	45.74	60.70	201.69
78	56.34	45.16	48.08	55.68	205.26
79	45.15	46.57	43.85	63.27	198.84
82	58.58	50.06	49.39	58.28	216.31
83	55.60	44.85	46.92	57.80	205.17
85	54.10	49.51	51.43	59.18	214.22
86	36.94	51.24	46.71	63.29	198.18
88	36.19	50.95	47.73	66.55	201.42
89	41.42	55.14	51.43	65.50	213.49
94	57.83	48.94	52.93	59.64	219.34
95	59.33	52.51	54.85	61.37	228.06
98	63.06	62.49	56.18	66.77	248.50
100	56.34	49.65	54.45	66.03	226.47
102	49.63	55.84	52.85	64.37	222.69
103	48.88	50.56	48.84	66.54	214.82
104	39.93	52.66	49.86	64.84	207.29
106	63.81	64.81	35.65	66.38	230.65
110	57.09	58.27	41.24	67.57	224.17
111	58.58	56.84	45.14	68.14	228.70
112	62.31	63.42	39.74	66.77	232.24
113	65.30	73.85	35.71	66.99	241.85
114	63.81	68.31	35.93	68.39	236.44
115	63.81	77.89	36.78	68.65	247.13
116	63.06	66.63	36.55	68.06	234.30
117	61.57	64.78	39.12	66.16	231.63
118	60.82	62.74	44.35	67.16	235.07
119	54.85	55.35	46.24	68.62	225.06
120	63.06	63.14	44.86	66.49	237.55
121	59.33	56.07	60.81	64.80	241.01
St. Paul					
14	39.87	54.88	40.16	64.01	198.92
15	40.51	47.87	40.09	53.55	182.02
30	39.87	47.63	39.54	58.57	185.61
32	48.73	47.47	51.36	61.24	208.80
33	60.76	49.68	55.21	59.68	225.33
43	36.08	18.19	38.82	56.55	149.64
48	70.25	36.00	70.59	69.64	246.48
49	64.56	49.80	61.26	62.27	237.89
50	63.39	56.96	60.30	59.12	239.77
53	62.03	56.06	61.47	61.34	240.90
56	62.66	50.58	57.63	59.73	230.60
57	65.19	55.72	63.29	60.74	244.94
58	63.39	46.69	61.02	57.52	228.62

Table A-1.2 concluded

CENSUS TRACT NUMBER	EDUCATION (Note a)	INCOME (Note b)	RENT (Note c)	POPULATION DENSITY (Note d)	TOTAL (Note e)
59	40.51	36.15	39.86	50.91	167.43
60	39.87	43.44	39.92	51.73	174.96
66	60.76	70.38	62.02	67.06	260.22
67	41.14	52.10	50.15	70.08	213.47
68	39.24	50.45	45.61	62.33	197.63
69	39.87	53.68	39.57	60.97	194.09
75	65.82	80.71	90.49	70.42	307.44

*In Minneapolis. For St. Paul computations based upon density per residential plus industrial acre.

**Computed from:
U.S. Bureau of the Census
U.S. Census of Population: 1950
Vol. III, Census Tract Statistics, Chapter 33
U.S. Government Printing Office,
Washington 25, D. C., 1952

Note a. Based upon median number of school years computed by persons twenty-five years and over, 1950

Note b. Based upon median income in dollars for families and unrelated individuals, 1949

Note c. Based upon median monthly contract rent per dwelling unit, 1950

Note d. For Minneapolis, based upon population density per residential plus commercial acre, 1950. For St. Paul, based upon population density per residential plus industrial acre, 1950. For both Minneapolis and St. Paul Z score used in determining socio-economic status obtained by subtracting computed Z score from 110.

Note e. Based upon the data used, the higher the total score, the higher the socio-economic status of the census tract population.

Table A-1.3

CLASSROOM SOCIO-ECONOMIC INDEX BASED UPON CENSUS TRACT DATA REGARDING MEDIAN INCOME, MEDIAN EDUCATION, POPULATION DENSITY, AND MEDIAN RENT FOR NORTH, SOUTH, SOUTHEAST, IN MINNEAPOLIS AND ST. PAUL GEOGRAPHICAL AREAS.

PLACEMENT BLOCK	SCHOOL WITH NUMBER AND TYPE OF CLASSROOMS USED	Z SCORE INDEX	PLACEMENT BLOCK	SCHOOL WITH NUMBER AND TYPE OF CLASSROOMS USED	Z SCORE INDEX
NORTH	5 (U)	203.43	SOUTHEAST	1 (L)	168.13
	9 (L)	220.85		2 (L)	222.07
	11 (L)	228.55		7 (L)	238.04
	12 (L)	207.82		(U)	238.04
	15 (U)	208.23		10 (L)	206.73
	19 (U)	194.76		(U)	206.73
				(U)	206.73
				13 (U)	165.87
	Area Mean	$\bar{X}=210.61$		14 (U)	231.59
	Area s.d.	12.19		16 (L)	221.10
SOUTH	3 (L)	238.31		(L)	221.10
	4 (L)	191.94		18 (U)*	232.53
	(L)	191.94			
	6 (L)	195.50			
	(L)	195.50		Area Mean	$\bar{X}=213.22$
	8 (L)	238.61		Area s.d.	24.39
	(U)	238.61	ST. PAUL	23 (L)	231.65
	17 (L)	228.17		24 (L)	283.83
	18 (U)*	232.53		25 (L)	192.55
	20 (U)	219.76		26 (L)	161.63
	21 (U)	205.09		(U)	161.63
	(U)	205.09		27 (U)	230.91
	22 (U)	239.66		28 (U)	192.27
	(U)	239.66		29 (U)	227.47
	Area Mean	$\bar{X}=218.60$		Area Mean	$\bar{X}=210.24$
	Area s.d.	30.33		Area s.d.	41.33

*NOTE: In order to meet the requirements of the experimental design pertaining to grade preference and assignment (upper and lower grades) one classroom in school #18 was put in the South placement block, the other classroom in the Southeast placement block.

Table A-1.4

ANALYSIS OF VARIANCE OF AREA SOCIO-ECONOMIC INDEXES

SV	df	SS	MS	F
Between Areas	3	484.74	161.58	0.2377 N.S.
Within Areas	36	24466.12	679.61	F.01 = 4.38
Total	39	24950.86		F.05 = 2.86

Table A-1.5

BARTLETT'S TEST OF THE HOMOGENEITY OF ESTIMATED VARIANCES OF AREA SOCIO-ECONOMIC INDEXES

Area	n(t)	s_t^2	V_t	$\log_e s_t^2$	$V_t \log_e s_t^2$	$\frac{1}{V_t}$
North 1	6	148.58 12.19	5	5.00112	25.00560	0.20000
South 2	14	919.93 30.33	13	6.82429	88.71577	0.07692
Southeast 3	12	595.13 24.39	11	6.38878	70.27658	0.09091
St. Paul 4	8	1708.53 41.33	7	7.43888	52.07216	0.51069
	40		36 = N		236.07011	0.51069

Appendix A-2.1

STUDENT TEACHING PROJECT
RATINGS OF CLASSROOM BEHAVIOR

Student Teacher ____________________
Observer ____________________
Date ____________________

Directions:

From your observations of the classroom behavior of this student teacher, rate each of the statements below using the following key. Your rating should reflect your judgment of the extent to which each statement would be generally true.

Key:

5. Always
4. Frequently
3. Occasionally
2. Seldom
1. Never

I. Creativity and Resourcefulness in Teaching

____ 1. Uses ingenuity in teaching-learning situations
____ 2. Uses children's experiences and examples appropriately
____ 3. Uses own experiences and examples appropriately
____ 4. Uses plans flexibly to reach anticipated goals
____ 5. Resolves unexpected situations resourcefully
____ 6. Recognizes children's learning problems
____ 7. Promotes the fluency of ideas in the classroom
____ 8. Adjusts level of communication to the group
____ 9. Helps children move from the concrete to the abstract (Builds toward generalizations)
____ 10. Helps children see relationships and develop syntheses
____ 11. Applies principles of child growth and development
____ 12. Draws upon knowledge of subject matter appropriately
____ 13. Uses teaching methods appropriately

II. Interpersonal Relations

____ 1. Uses a variety of teaching techniques to elicit and handle pupil response.
____ 2. Encourages interest and enthusiasm on the part of all children
____ 3. Assists pupils of all abilities to participate
____ 4. Encourages pupils to assume proper share in establishment of group goals
____ 5. Maintains businesslike yet accepting atmosphere
____ 6. Helps children work well together as a group
____ 7. Uses flexible and effective influence for pupil control
____ 8. Helps children assess progress toward individual and group goals

III. Estimate of Success

____ 1. From your observations what is your estimate of the student teacher's success?
    5. Like the best student teacher, 4. Between average and best, 3. As effective as the average, 2. Between average and poorest, 1. Similar to the poorest.

____ 2. From your observations what is your estimate of this student's probably success as a teacher?
    5. Will be like the best teacher, 4. Between average and best, 3. As effective as the average, 2. Between average and the poorest, 1. Similar to the poorest.

Appendix A-2.2

STUDENT TEACHING PROJECT
University Supervisor Ratings

STUDENT TEACHER ______________________

UNIVERSITY SUPERVISOR __________________

QUARTER: Fall Winter

DIRECTIONS: Circle the scale value that represents your judgment of the extent to which the student teacher is described by the statements at the right.

Scale 1

5 4 3 2 1

High Low

Student Teacher with high scale scores are characterized by

- sets definite structure and organization
- sets high standards
- provides rules, regulations, procedures
- interaction with pupils often characterized by
  1. persuasion, prohibition, commands
  2. blame, criticism, sarcasm
  3. threats, withholding approval
  4. lack of regard for student feeling
  5. intolerant of differences

Scale 2

5 4 3 2 1

High Low

Student Teacher with high scale scores are characterized by

- sets broad structure and organization
- provides for differentiation of tasks
- seeks common purpose in student differences
- flexibility of standards
- interaction with pupils characterized by
  1. support, encouragement, praise
  2. feeedom of student expression
  3. attention to student feelings
  4. encouragement of differences of approach
  5. promotion of acceptance of others

Appendix A-3
Observation Schedule For Oscar III

General Instructions

This observation sheet is designed to give a general picture of teacher and pupil activities and relationships; grouping; materials used, and subject content discussed during three different five minute observation periods. The sheet is divided into eight sections; each section is subdivided into groups and categories. Each category is defined and explained in detail in the section of this paper entitled "Definition and Explanation of Categories."

In using this observation sheet the observer is to check in the appropriate column, I, III, or V, for a period of five minutes duration. Since this observation sheet is, for this research project, to be used alternately with a different type observation sheet (Minnesota System), the observer marks this sheet during alternate five minute periods. It should be noted that the observer begins the observation period with this observation schedule.

In marking this observation sheet a check is made in the appropriate row and column for each different category item which occurs during each five minute observation period.

Note that any category on this observation sheet is checked only once for each five minute observation period regardless of how many times a particular activity or behavior occurs during the observation period.

Definition and Explanation of Categories

Section I

Group A: Teacher-pupil

1. Teacher questions, pupil answers. The Teacher's question may be either direct or implied. However, it must be a question for which an answer is expected, that is, the teacher must pause long enough for an appropriate answer to be given and the question cannot be rhetorical. This category is to be checked when a teacher calls on a pupil, points to a pupil, or in another way indicates that a pupil is to respond. A statement such as, "I don't know," or an incorrect answer, in this case, constitutes an answer. If the pupil ignores the question, category 4 of group K should be checked. This category is different than categories 1, 2, and 3 of group L in that, in this category, the questions require only brief answers or comments (thus different than L1 and L3) which are not dependent upon memorization by rote (thus different than L2).

Some examples of questions in this category are, "Billy, did you pass your paper to the front of the room," "Jane, which spelling word did you have the most difficulty learning," or, "Bob, do you understand how Jimmy worked this problem."

An answer to a question such as "Mary, tell the class why you liked story A better than story B," or, "In the history lesson you studied for homework, Tom, who was said to be the author of the Declaration of Independence," should be marked in categories 1 and 2 of group L, respectively.

2. Teacher answers pupil's question. As in A1, the question may be either direct or implied, however, it must be a question for which an answer is expected (See A1). This category is checked when, as indicated, the question is pupil

initiated and the teacher responds. If the teacher ignores the question, category 4 of group A should be checked.

3. Teacher interrupts pupil. This category is checked when the teacher speaks while a pupil is speaking regardless of the reason for the interruption, e.g., for clarification, additional information, change of topic, or correction.

4. Teacher ignores pupil's question. This category is checked whenever the teacher, either deliberately or inadvertently, ignores, or does not answer or comment on, a question by a pupil, provided that the pupil expects an answer to the question (see A1) and that the question is asked in such a way that it is, in all probability, audible to the teacher. This category is checked regardless of the occasion during which the question is asked, for example, during a test.

5. Teacher warns or threatens pupil or pupils. This must be a verbal threat or warning and the threat or warning must be explicit in the statement of the teacher. A threat or warning would ordinarily take the form, "If you do (don't) do X then Y (will happen)," or, "If A does (doesn't) do X then Y (will happen)." More specifically, "If you don't get the problems finished before class is dismissed you will stay after school," or, "If Billy doesn't stop misbehaving he will have to leave the room."

6. Teacher uses criticism and/or sarcasm. This category includes scorn, ridicule, or (contemptuous) laughter, but the statement or the laughter must be such that the content, the context, or the manner of expression explicitly indicates or portrays criticism, sarcasm, scorn, or ridicule. Academic correction, e.g., three times four is twelve not nine are not included in this category unless criticism for not knowing the correct answer, or sarcasm, accompanies the correction.

7. Teacher praises and/or encourages pupil or pupils for good work. The praise or the encouragement may take the form of a comment or a remark to the individual concerned, e.g., "This is the best arithmetic paper you have submitted this week;" to the class about the individual, e.g., "If the rest of the class did their problems as well as X.....;" to a specified part of the class, e.g., "Those of you who solved problem five correctly have learned the multiplication tables well," or to the entire class, e.g., "All of you must have studied very hard to have learned the multiplication tables so well in such a short time." This category is checked when the praise or the encouragement is for a task or an activity well done.

8. Teacher praises and/or encourages pupil or pupils for good behavior. This category is the same as A7 except that the praise or encouragement is for good behavior rather than for good work. Thus, the praise or encouragement may take the form of a comment or a remark to the individual concerned, to the class about the individual, to a specified part of the class, or, to the entire class.

9. Teacher works with individual pupil. This category is checked when the teacher aids an individual pupil regardless of whether the aid is asked for by the pupil and regardless of whether the aid is of an "academic" nature or not (it might, for example, be helping a pupil put on boots or fix a toy).

Group B: Teacher-class

1. Teacher lectures. This category is checked during the time when the class is conducted in a manner which makes no provision for active pupil participation in the lesson being presented except through some form of intrusion or interruption in the teacher's presentation. If an interruption occurs or should

active participation on the part of the pupils be initiated then another category is to be checked (see A2, A4, or B3).

2. Teacher reads, tells story. This category is checked whether the story is told or the passage read by the teacher to clarify or to further elaborate a particular point in a lecture or whether the story or the reading itself is the primary concern of the teacher and the class at the particular time.

3. Teacher talks to class. This category is checked when the teacher is discussing topics of a non-academic nature, for example, running in the hall, good sportsmanship, obeying the school patrol; if the teacher is giving directions to the class, e.g., "Turn to page 41 of your arithmetic books," or when the teacher is answering or asking questions so that there is intermittent pupil participation during the imparting of information.

4. Teacher illustrates at board. This category is checked when the teacher uses the blackboard as a means of facilitating explanation, elaboration, clarification, justification, or the imparting of information.

5. Teacher illustrates with map or chart. This category includes the use by the teacher of graphs, pictures, posters, or other pictorial material and is checked when one of these visual aids is used by the teacher as a means of facilitating explanation, elaboration, clarification, justification, or the imparting of information.

6. Teacher demonstrates. This category is checked when the teacher uses a three-dimensional model, object, or specimen as a means of facilitating explanation, elaboration, clarification, justification, or the imparting of information.

7. Teacher shows film, slides, plays records. This category is checked when the teacher for any reason whatsoever, utilizes these particular audio, visual, or audio-visual materials.

8. Teacher passes books, paper, milk. This category is checked whenever the teacher is involved in the distribution or the passing of a book or books, paper, or milk.

9. Teacher leads singing, exercise, or games. This category is checked when the teacher is actively engaged as the director of singing drill exercise, or games.

Group C: Teacher Activities

1. Teacher works at desk (or at a temporary location serving as a desk.) This category is checked at those times when the teacher is engaged in making out reports, grading papers, composing lesson plans, and similar tasks. Should the teacher temporarily stop this activity in order to help a pupil or pupils, a check should be made in A9 or in some category of Group W or X.

2. Teacher cleans, decorates room. This category is checked when the teacher is actively engaged in cleaning (picking up paper, dusting) or decorating the room.

3. Teacher writes on, decorates board. This category is checked whenever the teacher, while the class is busy at some other task, decorates or writes on or posts information on a bulletin board or blackboard. This category is distinct from, and is not to be confused with category 4 of Group B.

4. Teacher leaves room. This category is checked when the teacher leaves the room except for the time when the teacher stands in the doorway or in the hall within sight and hearing of the class.

Section II

Group K: Pupil-teacher

1. Pupil whispers. This category is checked only when the whispering on the part of the pupil indicates or reflects, by the time and the place of its occurrence, inattention on the part of a pupil or pupils. Thus, this category is checked only when the whispering occurs while the teacher is expecting attention.

2. Pupil laughs. This category is checked only when the laughter is audible and when the laughter indicates or reflects, by the time and the place of its occurrence, inattention on the part of the pupil or pupils. Laughter with the class at an amusing story would, for example, not be included in this category while laughter resulting from the telling of a "private joke" would be included in this category.

3. Pupil interrupts teacher. This category is checked when the pupil speaks while the teacher is speaking regardless of the reason for the interruption, e.g., for clarification, additional information, change of topic, or correction.

4. Pupil ignores teacher's question. This category is checked whenever the pupil, either deliberately or inadvertently, ignores, or does not answer or comment on, a question of the teacher provided that the teacher expects an answer to the question (see A1) and that the question is asked in such a way that it is, in all probability, audible to the pupil. This category is checked regardless of the occasion during which the question is asked, for example, during a test or a play period.

5. Pupil shows hostility toward teacher. This behavior may be either verbal or non-verbal, but it must be such as to clearly indicate, not merely imply, hostility. The hostility may be in the form of a direct refusal, a definite resistance, or a distinct reluctance on the part of the pupil to perform some task or to take some particular course of action; it may be a specific comment of a derogatory or derisive nature made to the teacher or to another pupil about the teacher, or it may take the form of an expression or gesture made by the pupil to the teacher or behind the teacher's back, e.g., to stick out his tongue or to "make a face" at the teacher. This category would also be checked should the pupil hit, or strike at, or throw some object at the teacher.

6. Pupil shows affection for teacher. This behavior may be either verbal or non-verbal, but it must be such as to clearly indicate, not merely imply, affection for the teacher. The affection may be expressed in the form of a comment or remark to the teacher or to another pupil about the teacher, e.g., "You are my favorite teacher," "You are nice," "I like Miss X," or it may be a direct expression of affection, i.e., putting his arm around the teacher or taking the teacher's hand.

Group L: Pupil-class

1. Pupil talks to group. This category is checked when a pupil talks

extemporaneously, except for those times when the pupil simply gives brief answers to question (see A1). This category is checked when, for example, a pupil tells what he saw on the way to school, what he liked best at the fair, or what he thought of a story.

2. Pupil recites. This category is checked when the pupil presents material learned by rote, usually in the form of a memorized poem or story or a particular fact or set of facts such as, for example, "What state borders Minnesota on the east," "Name the thirteen colonies," "Who invented the cotton gin."

3. Pupil reports, gives prepared talk. This category is checked when the pupil gives an organized and prepared report or talk whether from memory (not to be confused with L1), read from a paper; or given with the aid of note or reference cards.

4. Pupil reads aloud. This category is checked when the pupil, usually at the request of the teacher, reads aloud from the text, a workbook, or similar material.

5. Pupil demonstrates, illustrates. This category is checked whenever the pupil uses the blackboard, a three-dimensional model, a map, chart, graph, picture, poster, or other pictorial material as a means of facilitating explanation, elaboration, clarification, or the imparting of information. Note that this category is different and distinct from M4.

6. Pupil gives skit or play. This category is checked when the pupil takes an active part in the presentation of a skit or a play.

7. Pupil sings, plays instrument. This category is checked whenever a pupil sings or plays an instrument whether this is done alone or in a group.

8. Pupil plays game. This category is checked whenever two or more pupils play a "recreational" or an "educational" game. If only one person is involved, such as working a puzzle, M2 is checked.

9. Pupil leads class. This category is checked whenever a pupil assumes the responsibility for a class presentation or a class function of an academic nature, i.e., excluding singing, exercise, games (see L10). In this category, the pupil ordinarily operates as a class chairman, class president, panel moderator or member, or in a similar capacity. This category is different and distinct from L1, L2, L3, and L4 for in these categories the teacher remains the active class leader despite the pupil contribution.

10. Pupil leads singing, exercise, games. This category is checked whenever a pupil leads the class in singing, exercises, or games.

11. Pupil passes book(s), paper, milk. This category is checked whenever a pupil, or pupils, with or without the consent or the aid of the teacher, performs the specified or similar activities.

Group M: Pupil activities

1. Pupil reads, studies at seat. This category is checked whenever the pupil reads or studies at his desk, or at a location which is being temporarily used as a desk. If pupil is both reading and writing, M2 is checked.

2. Pupil writes, manipulates at seat. This category is checked whenever a

pupil does any writing or calculating or manipulating of some object, for example, an abacus, clock model, or flash card at his desk, or at a location which is being temporarily used as a desk. This is, however, not checked when the pupil is doing art or craft work (see M3).

3. Pupil paints, draws, cuts, pastes at seat. This category is checked whenever a pupil is engaged in some type or kind of art or craft work, either at his desk or at some location used by the pupil as a temporary desk.

4. Pupil works at board. This category is checked whenever a pupil uses the blackboard, either at his own discretion or at the direction of the teacher, except when the pupil uses the blackboard as a means of facilitating explanation (see L5) or when the pupil decorates the board (see M5).

5. Pupil decorates room, board. This category is checked whenever a pupil, with or without the direction or permission of the teacher, decorates a bulletin board, blackboard, window, or wall.

6. Pupil cleans room, board. This category is checked whenever a pupil, with or without the direction or request of the teacher, erases or washes the blackboard, picks up paper from the floor, or performs similar tasks.

7. Pupil rests, has snack. This category is checked whenever a pupil, with or without permission of the teacher, rests on a mat, or at his seat, or has a snack.

8. Pupil leaves room. This category is checked whenever a pupil, with or without permission of the teacher, leaves the classroom.

Group N: Pupil-pupil

1. Pupil scuffles, fights with pupil. This category is checked whenever two or more pupils engage in hitting, slapping, wrestling, or some other form of scuffling or fighting. This category is thus checked only when force or contact is present.

2. Pupil shows hostility toward pupil. This category is checked whenever a pupil throws some object, for example, paper, eraser, pencil, or paper clip at another pupil; when a pupil directly threatens or warns another pupil, i.e., "If X (behavior) continues or occurs then Y will happen (as a result of his action), "or when the threat or warning is explicit in the action of the pupil, e.g., shaking a fist at a pupil; or whenever a pupil uses words which, in their context, indicate anger, bitterness, scorn, ridicule, sarcasm, rebuff, or derision of, or for, toward, or about, another pupil. Thus, the hostility may be either verbal, or non-verbal, or both. It may be direct or indirect, i.e., it may be to or about the pupil concerned. However, it must be clearly indicated hostility by the action, the expression, the context, or by all three.

3. Pupil shows affection toward pupil. This category is checked when a pupil talks or laughs with another pupil, when he takes another pupil's hand, puts his arm around another pupil, smiles at another pupil, or makes some remark such as, "I like you," or "I like X" to or about another pupil. Thus, the expression of affection may be either verbal or non-verbal or both. However, it must be a clear expression of affection by the action, the expression, the context, or by all three.

Section III

Group V: Grouping

1. Teacher choice of volunteers. This category is checked when the teacher makes the choice of persons to be part of a particular group, or to perform some activity or task entirely from among those pupils who volunteered (by raising their hand or some other method) for the task, or activity, or group. Note: If the selection is based in part on volunteers and in part on some other basis, V2 should be checked.

2. Teacher choice on some arbitrary basis. This category is checked when the teacher makes the choice of person's to be part of a particular group, or to perform some activity or task, partially or entirely on some basis other than pupils who volunteered, e.g.,"All those that missed problem three come to the front of the room." If it is not certain whether the pupils volunteered for the particular group, or activity, or task, for example, "Now that we have finished the spelling lesson you may re-form your social studies committees and continue to work on your special group problems," this category is to be checked.

3. Pupil choice of volunteers. This category is checked when a pupil, as a class or group leader, is responsible for and makes the selection of other pupils for a particular group, or activity, or task, entirely from among those pupils who volunteered (by raising their hands or by some other method) for the group, or activity, or task.

4. Pupil choice on some arbitrary basis. This category is checked when a pupil, as a class or group leader, is responsible for and makes the selection of other pupils for a particular group, or activity, or task, in whole, or in part, on some basis other than volunteers.

5. Pupil goes to the group of his choice. This category is checked whenever the pupils are free to go to any group of their own choosing, for example, "All of you who wish to help with a report on George Washington go to the table in the back of the room, those of you who wish to help with a report on Abraham Lincoln can use my desk, and those of you who wish to help make a report on Woodrow Wilson can use the table at the side of the room."

Group W:

Individual task with teacher. A task is an individual task when a pupil is engaged in some activity or project in which the pupil alone is responsible for determining and/or accomplishing a goal and/or the means for attaining that goal. "With teacher" means that the teacher is not merely present in the room but that he is engaged in aiding an individual pupil or pupils. W1, W2, or W3 is checked depending upon the size of the group, i.e., over one half the class, four pupils to half the class, or two or three pupils.

Group X:

Group task with teacher. A task is a group task when two or more pupils are engaged in some activity or project in which the pupils together--through discussion and planning--are responsible for determining and/or accomplishing the goal and/or the means for attaining that goal. It should be noted that a category in this group is not checked when two or more pupils are performing the same task individually. The remainder of the distinctions made in this group are the same as in Group W.

Group Y:

Individual task without teacher. "Without teacher" means that the teacher is not engaged in aiding an individual pupil or pupils. It should be noted that some category in Groups Y or Z (Groups "without teacher") would be checked when, for example, the teacher is in the classroom but is grading papers or doing some other "desk" work.

The remainder of the distinctions which are necessary to make regarding the categories of Groups Y and Z have been discussed in connection with Groups W and X.

Sections IV and V

These sections pertain to the materials used by the teacher and the pupils respectively in carrying out their activities. The materials included are:

1. blackboard
2. map, chart, or picture
3. slide, film strip, or film
4. audio aid (such as tape recorder, microphone, or record)
5. three dimensional object (such as model)
6. textbook or workbook
7. supplementary reading materials
8. writing materials
9. handicraft or art materials
10. special teaching or learning materials not itemized
11. no material used

If special teaching or learning materials which are not itemized are used by the teacher and/or the pupils, the observer should specify at the bottom of the appropriate group the particular material or materials being used.

Section VI

This section pertains to the content area or areas which are discussed by the teacher and the pupils during the observation period. The content area on which the teacher spends the majority of his time during the observation period should be checked twice. If "test" is checked, some other category or categories pertaining to the subject area or areas the pupils are being tested on should also be checked.

Section VII

This section indicates whether or not the supervising teacher was present in the room during the observation period (1), and, if present, whether he made any comments to the class or to the student teacher during the observation period (2).

Appendix A-4

The Minnesota System of Categorizing Verbal Behaviors

The Minnesota System was used by the research project personnel in order to categorize the verbal behavior of the student teachers and the pupils in their classrooms. This observation schedule was used during three alternate five-minute periods--the Oscar III was used during the other three five-minute observation periods--beginning with observation period number two. A slight modification was made in the Minnesota System in order to better adapt it for use in this research project.

Categories were adapted from the Minnesota System of Analysis in which verbal behaviors are classified in sequence. The categories, according to Flanders:

> represent different qualities of action and are based on a theory of teacher influence. The teacher can be direct and restrict the freedom of action of a student, the teacher can be indirect and increase the freedom of action of students, or he can, within a single contact, set a particular balance of indirect-direct influence. The categories are assumed to be totally inclusive of communication events. The procedure of tallying in one, and in only one, category makes the categories mutually exclusive.

The modification in the Minnesota System as it was used in the student teacher research project occurred in the frequency of making tally marks. For the purposes of the research project it was deemed advisable to tally every category change, or, if there was no change in category, to make a tally approximately every three seconds. This was done rather than tallying at approximately the three second intervals only. In order to insure the recording of maximum interaction between the student teacher and the pupils it was further decided that the research project observer would stop recording tallies if the student teacher stopped teaching, for example, during the showing of a movie or during extended periods of silence or of confusion. In cases in which the observer stopped recording tallies the observation was continued as soon as possible and the break in the observation period was recorded along with the time duration of the break.

The research project observers used mimeographed tally sheets for recording the tallies for the three observation periods. In addition to totaling the tallies for the three observation periods, the rate was computed for each category. The rate was computed by dividing the tallies in each category for each observation period by the total tallies for the particular observation period.

The Minnesota System is designed to categorize all verbal classroom behaviors. However, the schedule is primarily designed to differentiate among and to categorize different kinds of teacher verbal behavior, and to give some indication of how the teacher's verbal behavior affects the verbal behavior of the pupils. Thus, categories one through nine are used to classify different kinds of teacher statements while categories ten and eleven are used to differentiate between self-initiated or spontaneous pupil response and teacher-directed pupil response. Category twelve is used to indicate either silence, for example, during a test or when pupils are reading quietly at their desks, or confusion, for example, when pupils are moving from their desks to form special work or reading groups, when several pupils are talking at the same time, or when there is laughter.

The following twelve categories are included in the Minnesota System. These categories are separate and distinct, and no scale value is implied by the number.

Category one pertains to the teacher's reaction to pupil feeling. In order for an entry to be made in this category the teacher must either:

a. respond to the emotional tone of a pupil's behavior in a positive and non-threatening manner

b. accept, clarify, or interpret to help the pupil relate to a problem or to the group.

Examples of the type of teacher statement which would be classified in this category are, "You feel that this assignment is too difficult for you," "You are afraid that if you take your library book home it will become lost."

Category two includes teacher praise or encourgement. The following kinds of teacher behavior are included in this category:

a. praising the work of a pupil or pupils, complimenting a pupil or pupils.

b. individual or group encourgement

c. making jokes or providing other forms of tension release in order to improve the classroom atmosphere without sacrificing any pupil or pupils.

d. nodding, "yesing," or "um-humming" to keep a pupil participating.

Category three includes teacher reaction to student ideas. This category is much the same as category one except that category three pertains to pupil ideas rather than to pupil feelings. The following are included in category three:

a. teacher response to the ideational aspects of a pupil's behavior in a positive and non-threatening manner

b. teacher summary of a pupil's response or comment (this can also include writing the pupil's contribution on the blackboard.

It can be seen from this description that in categories one, two or three the orientation of the teacher is primarily toward the pupil: toward the pupil's ideas or feelings.

Category four is used when the teacher asks a question. An entry would be made in this category if the teacher:

a. asks about class procedure

b. asks about the content of the material being discussed (asks if there is need for additional clarification, explanation, or elaboration)

c. calls on a pupil who has his hand raised or otherwise indicates that he wishes to speak.

Category five pertains to administrative or routine statements by the teacher. This category is used when:

a. the teacher makes comments or requests not directly concerned with the immediate learning activity, for example, asking a pupil to close the door or window, or to take the attendance or some other report to the office

b. talks with a person, or persons, who is not a member of the class, for example, a pupil from another class, the school principal, or school nurse.

An entry is made in category six when the teacher gives information or an opinion. This category is used when the teacher:

a. lectures

b. states opinions or facts

c. asks categorical questions.

Category seven is used when the teacher gives directions. An entry is made in this category when the teacher:

a. makes requests or gives commands pertaining directly to the learning activities such as, "Turn to page 734 of your reading text," or "Tomorrow work all the problems on page 42 or your arithmetic workbook."

b. calls on a pupil to recite, to answer a question, or otherwise to respond without the pupil having indicated, by raising his hand or in some other manner, a desire to do so.

Category eight pertains to teacher criticism. The criticism is intended to change present pupil behavior or attitudes and may be directed toward an individual pupil, a group of pupils, or the entire class.

Category nine is used to classify teacher self-justification. This category is used when the teacher:

a. makes a remark or remarks explaining the reason for her behavior

b. justifies her control of a pupil, or pupils, or of the entire class

c. uses self-reference or relates anecdotes about herself which are not directly related to the class activities.

Category ten is used for self-initiated pupil talk. An entry is made in this category when:

a. the pupil "speaks out" in class

b. the pupil is called upon by the teacher after indicating, by raising his hand or in some other manner, that he wishes to speak.

c. after being called upon by the teacher, without originally indicating a desire to respond, the pupil goes on to enlarge upon, expand, qualify, or illustrate his initial response. In this case the original entry or entries would be made in category eleven followed by an entry or entries in category ten.

Category eleven pertains to directed pupil talk. An entry is made in this category when:

a. the teacher calls on the pupil for a response although the pupil has given no indication of a desire to speak.

b. the pupil reads a prepared paper or makes a previously prepared class presentation

c. the class sings, or reads, in unison.

Category twelve is used to indicate silence or confusion. This category is used when:

a. it is not clear who is talking, i.e., several people are talking at once.

b. there is laughter

c. the pupils are moving from one location to another, such as from their desks to reading groups

d. there is silence, for example, during a test or during the time pupils read quietly at their desks.

It can be seen from the previous discussion that an entry can change from one category to another during either the teacher's or the pupil's remarks. Hence, the entry for pupil response can, as has been indicated, change during a response from category eleven to category ten. The teachers's comments can, of course, be classified in several different categories during a brief discourse. Thus, the statements of the teacher may, within a short interval of time, go from summarizing the response of a pupil (category three), to praising the pupil's response (category two), to enlarging upon, or presenting in more detail, the response of the pupil (category six), to asking a question about the further elaboration (category four), to calling upon a pupil for a response to the question asked (category seven). If the teacher paused in the above sequence to "correct" a pupil and to explain why she found it necessary to correct him, an entry, or entries, would also have been made in categories eight and nine.

Appendix B

Frequency of Occurrence of Classroom Behaviors Recorded in Ten Observations of Thirty-six Elementary Student Teachers Using the Categories of the OScAR III

			IBM				Frequency				
Category			Col	0	1	2	3	4	5	6	7
A.	1.	t qu, p ans	1	4	33	110	213				
	2.	t ans p qu	2	65	94	128	73				
	3.	t intrrpts p	3	241	81	20	18				
	4.	t ignrs p qu	4	299	56	4	1				
	5.	t wrns, thrtns p	5	281	55	19	5				
	6.	t criz srcsm	6	78	92	113	77				
	7.	t prs pr gd wk	7	138	102	83	37				
	8.	t prs p go bhv	8	319	37	4	0				
	9.	t wks w ind p	9	144	108	81	27				
B.	1.	t lctrs	10	173	85	66	36				
	2.	t rds, tls sty	11	311	40	8	1				
	3.	t tks t cl	12	15	49	98	198				
	4.	t ills at bd	13	177	98	69	16				
	5.	t ills wth mp, cht	14	284	44	21	11				
	6.	t dmnstrs	15	310	35	12	3				
	7.	t shws flm, sld, pls rcd	16	342	14	0	4				
	8.	t pss bk, ppr, mlk	17	291	57	9	3				
	9.	t lds sng, ex, gm	18	287	50	14	9				
c.	1.	t wks at dsk									
	2.	t clns, dcts rm									
	3.	t wrts on, dcts bd	19	242	78	24	14	0	1	0	1
	4.	t lvs rm									
	5.	t tlks wth vstr									
k.	1.	p whsprs	20	166	84	116	94				
	2.	p lghs	21	241	73	29	17				
	3.	p intrrpts t	22	200	92	46	22				
	4.	p igns q of t	23	316	35	8	1				
	5.	p shws hst t t	24	316	34	5	5				
	6.	p shws af t t	25	340	18	1	1				
L.	1.	p tks t gp	26	140	100	72	48				
	2.	p rcts	27	84	109	88	79				
	3.	p rpts, gvs prpd tk	28	339	15	6	0				
	4.	p rds ald	29	169	89	70	32				
	5.	p dmnstrs, illus	30	271	67	19	3				
	6.	p gvs skt, ply	31	351	7	2	0				
	7.	p sngs, pl instr	32	326	19	7	8				
	8.	p pls gm	33	340	16	4	0				
	9.	p lds cls	34	330	23	6	1				
	10.	p lds sng, ex, gm	35	344	14	1	1				
	11.	p pss bks, ppr, mlk	36	274	72	14	0				

Appendix B con't.

Category			IBM				Frequency				
			Col	0	1	2	3	4	5	6	7
M.	1.	p rds, stdys at st	37	133	70	64	39	25	15	14	0
	2.	p wrts, mnps at st									
	3.	p pnts, cts, drws									
	4.	p wks at bd	38	283	48	20	7	1	0	1	0
	5.	p dcts rm, bd									
	6.	p clns rm, bd	39	315	36	6	2	0	0	1	0
N.	1.	p scffls, fghts with p	40	323	25	12	0				
	2.	p shws hst to p	41	267	62	27	4				
	3.	p shws aff t p	42	329	27	4	0				
V.	1.	t chs vlntrs									
	2.	t chs	43	318	19	12	11	0	0	0	0
	3.	p chs vlntrs									
	4.	p chs	44	356	2	2	0	0	0	0	0
	5.	p t gp of chc									
W.	ind tsk w/t										
	1.	over 1/2 cl wt									
	2.	4-1/2 cl wt	45	330	16	6	8	0	0	0	0
	3.	2-3 wt									
X.	gp tsk w/t										
	1.	ovr 1/2 cl wt									
	2.	4-1/2 cl wt	46	335	11	9	5	0	0	0	0
	3.	2-3 wt									
Y.	ind tsk w/o t:										
	1.	over 1/2 cl w/ot									
	2.	4-1/2 cl w/ot	47	336	10	9	5	0	0	0	0
	3.	2-3 w/ot									
Z.	gp tsk w/o t										
	1.	ovr 1/] cl w/o t									
	2.	4-1/2 cl w/o t	48	357	0	2	0	0	1	0	0
	3.	2-3 w/o t									

Appendix B - concluded

Category	IBM Col	Frequency 0	1	2	3	4	5	6	7
Teacher Materials									
1. blbd									
2. mp, cht, pctr									
3. sld, fm, etc									
4. audio aid	49	68	86	86	85	22	10	2	1
5. 3 dim object									
9. hcft, art									
10. spec tchg aid*									
6. txt, wkbk									
7. supl rdg mtr	50	88	64	99	97	7	4	1	0
8. wrtng									
Pupil Materials									
1. blbd									
2. mp, cht, pctr									
3. sld, fm, etc									
4. audio aid	51	151	99	57	40	5	3	3	2
5. 3 dim object									
9. hcft, art									
10. spc lrng aid**									
6. txt, wkbk									
7. supl rdg mtr	52	50	44	78	95	36	25	24	8
8. wrtng									
1. rdg	53	197	41	36	87				
2. math	54	284	31	22	22				
3. lang arts	55	277	35	32	16				
4. soc st	56	235	31	36	58				
5. science	57	336	4	7	13				
6. recreation	58	315	32	8	5				
7. art, craft	59	323	25	4	8				
8. music	60	333	9	5	13				

Appendix C

Table C-1 - Regression analysis predicting n Dom(RA) scores.

A Matrix

	x_1 (HSR)	x_2 (GPA)	x_3 (Miller)	x_4 (MTAI)	x_5 (Bowers)	x_6 (Counselor) N-Dom
x_1	5430.2222	104.9800	1084.6667	-354.5556	-219.8889	42.4444
x_2	104.9800	7.6400	58.9250	-2.6750	-6.0000	2.6250
x_3	1084.6667	58.9250	2839.0000	3032.3333	102.3333	15.3333
x_4	-354.5556	-2.6750	3032.3333	25687.8889	-1022.7778	34.8889
x_5	-219.8889	-6.0000	102.3333	-1022.7778	2243.5556	-15.7778
x_6	42.4444	2.6250	15.3333	34.8889	15.7778	29.8889

Table C-1 con't.

A^{-1} Matrix

x	x_1 (HSR)	x_2 (GPA)	x_3 (Miller)	x_4 (MTAI)	x_5 (Bowers)	x_6 (Counselor) N-Dom
x_1	.0002553283	-.0031356143	-.0000427804	.0000092245	.0000223436	-.0000642292
x_2	-.0031356143	.2054469260	-.0034732004	.0004247428	.0005095244	-.0120333718
x_3	-.0000427804	-.0034732006	.0005105968	-.0000640071	-.0000649353	.0001442754
x_4	.0000092245	.0004247428	-.0000640071	.0000477850	.0000263254	-.0000594477
x_5	.0000223437	.0005095250	-.0000649353	.0000263254	.0004654447	.0001718041
x_6	-.0000642292	-.0120353715	.0001442750	-.0000594477	.0001718059	.034691508

Table C-1 con't.

G Matrix	B^1 Matrix
$g_1 = -52.45555$	$b_1 = -.00058468$
$g_2 = -4.59463$	$b_2 = -.82926519$
$g_3 = -16.91666$	$b_3 = +.02494831$
$g_4 = -232.76111$	$b_4 = -.01267242$
$g_5 = -6.27777$	$b_5 = -.01137154$
$g_6 = +.53888$	$b_6 = +.08767972$

$B^1G = +6.48707766$

$R^2 = .09467922$

$R = .307699$

$F = .5054$ n.s.

Table C-2 - Regression analysis predicting n Intg(RA) scores.

A Matrix

	x_1 (HSR)	x_2 (GPA)	x_3 (Miller)	x_4 (MTAI)	x_5 (Bowers)	x_6 (Counselor) N-Intg
x_1	5430.2222	104.9800	1084.6667	-354.5556	-219.8889	-4.5556
x_2	104.9800	7.6400	58.9250	-2.6750	-6.0000	2.2150
x_3	1084.6667	58.9250	2839.0000	3032.3333	102.3333	-47.6667
x_4	-354.5556	-2.6750	3032.3333	25687.8889	-1022.7778	95.8889
x_5	-219.8889	-6.0000	102.3333	-1022.7778	2243.5556	-18.7778
x_6	-4.5556	2.2150	-47.6667	95.8889	-18.7778	15.8889

Table C-2 con't.

A^{-1} Matrix

	x_1 (HSR)	x_2 (GPA)	x_3 (Miller)	x_4 (MTAI)	x_5 (Bowers)	x_6 (Counselor) N-Intg.
x_1	.0002579281	-.0034695689	-.0000250123	.0000053036	.0000235635	.0004784338
x_2	-.0034695696	.2370013224	-.0054294396	.0008409912	.0004657517	-.0548472132
x_3	-.0000250123	-.0054294413	.0006226538	-.0000882910	-.0000598450	.0030797734
x_4	.0000053036	.0008409915	-.0000882910	.0000530249	.0000253558	-.0006706240
x_5	.0000235635	.0004657518	-.0000598450	.0000253558	.0004648929	.0001586888
x_6	.0004784347	-.0548471837	.0030797691	-.0006706232	.0001586888	.0841935104

Table C-2 con't.

G Matrix	B^1 Matrix
g_1 = 81.20000	b_1 = -.0113471298
g_2 = 10.32425	b_2 = 1.4904832012
g_3 = 163.95000	b_3 = .0024708548
g_4 = 694.95000	b_4 = .0274936243
g_5 = 26.30000	b_5 = .0278608901
g_6 = 6.95000	b_6 = .1007898251

B^1G = 35.4117554348

R^2 = .358463929

R = .598719

F = 2.700 Sig. at .05 level